UNCLASSIFIED

AD 208613

Armed Services Technical Information Agency

ARLINGTON HALL STATION
ARLINGTON 12 VIRGINIA

FOR MICRO-CARD CONTROL ONLY	1 OF 2

NOTICE: WHEN GOVERNMENT OR OTHER DRAWINGS, SPECIFICATIONS OR OTHER DATA ARE USED FOR ANY PURPOSE OTHER THAN IN CONNECTION WITH A DEFINITELY RELATED GOVERNMENT PROCUREMENT OPERATION, THE U. S. GOVERNMENT THEREBY INCURS NO RESPONSIBILITY, NOR ANY OBLIGATION WHATSOEVER; AND THE FACT THAT THE GOVERNMENT MAY HAVE FORMULATED, FURNISHED, OR IN ANY WAY SUPPLIED THE SAID DRAWINGS, SPECIFICATIONS, OR OTHER DATA IS NOT TO BE REGARDED BY IMPLICATION OR OTHERWISE AS IN ANY MANNER LICENSING THE HOLDER OR ANY OTHER PERSON OR CORPORATION, OR CONVEYING ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE OR SELL ANY PATENTED INVENTION THAT MAY IN ANY WAY BE RELATED THERETO.

UNCLASSIFIED

208613

FILE COPY

2

Return to

ASTIA

ARLINGTON HALL STATION

ARLINGTON 12, VIRGINIA

Attn: TISSS

BEST

AVAILABLE

COPY

Statistical Techniques Research Group
Section of Mathematical Statistics
Department of Mathematics
Princeton Univeratty, Princeton, N.J.

Technical Report No. 14

PERMUTATION SUPPORT FOR MULTIVARIATE TECHNIQUES

by

Harvey J. Arnold

Department of Army Project No. 5B99-01-004
Ordnance R and D Project No. PB2-0001
OOR Contract No. 1715
Contract No. DA 36-034-ORD- 2297

ABSTRACT

Many writers have been concerned with the problem of justifying the use of standard tests based on normal theory assumptions when in fact the underlying distribution is suspected of being non-normal. This research has been concerned mainly with tests based only on univariate situations. They have found that the standard t and F tests are remarkably insensitive to deviations from normality.

Here, the multivariate generalization of the t test is investigated. The first four permutation cumulants are determined for a statistic which is a simple function of the commonly used Hotelling's T^2. In the univariate situation T^2 simply becomes the square of t, and tests based on T^2 are identical with tests based on t. Complete details of the investigation are carried through for only two dimensions but similar methods should apply to more dimensions. The derived permutation cumulants are applied to data obtained from a sampling experiment. The samples are from bivariate normal and rectangular populations. The samples drawn from the normal population served as a verification of the theoretical computations and also provided a comparison for the empirical results based on the samples from the rectangular population. The actual empirical calculations consisted in obtaining all the permutations of the statistic used (a function of T^2) for each sample and obtaining frequency distributions of the number of permutations which give values for the statistic which are greater than those obtained from normal theory assumptions for a certain significance level. It is observed that

there is mild disagreement between the nominal percentage point of the test based on the assumption that the underlying distribution is normal and the actual percentage point obtained by considering all permutations of each sample from the rectangular population.

A discussion of various ways of adjusting the test criterion is included when one suspects that the data does not come from a normal distribution. These methods are used on the samples from the rectangular distribution. One of these methods provides excellent agreement at the various significance levels investigated.

Suggestions toward directions for future investigations are given.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION AND ONE-DIMENSIONAL CASE

1.1 Introduction

The problem of justifying or modifying analysis of variance procedures when the form of the population governing the observations is not known to be normal has been investigated by Pitman [12], Welch [14] and others [1, 3, 7, 8, 9]. The earlier papers of Pitman [10, 11] made it clear that he originally set out to find an approximately non-parametric test based on quadratic expressions. Pitman [12] goes deeper into the randomization theory for the simple two-way layout than does Welch.

The problem considered here is the extension of Pitman's analysis-of-variance randomization test to the case of more than one response. Such an extension might, in the case of two treatments, be expected to lead to justification or modification of the use of Hotelling's T^2 [5]. Full details have been carried through for only the case of two treatments and two responses.

2. Pitman's Treatment of the Univariate Case

Since the multivariate problem will be treated along the lines of Pitman's treatment of the univariate case, it seems logical to begin by outlining his procedure. He considers k blocks (called "batches" by Pitman), each consisting of n individuals. Each set of n individuals is not necessarily considered as having come from a larger population.

The n treatments are assigned to the individuals of a block at random. The experimental arrangement is thus in the form of randomized blocks and the interest is in testing whether differences in treatments have produced any real differences in the response measured.

For k blocks, each with n treatments, the responses x_{ij} can be displayed as follows

$$\begin{array}{cccc} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{array}$$

where $x_{1r}, x_{2r}, \ldots, x_{kr}$ are the responses of the k individuals subjected to treatment r.

Let

$$x_{i.} = \frac{1}{n}\sum_{j=1}^{n} x_{ij} ,$$

$$x_{.j} = \frac{1}{k}\sum_{i=1}^{k} x_{ij} , \qquad 1.1$$

and $$x_{..} = \frac{1}{n}\sum_{j=1}^{n} x_{.j} = \frac{1}{k}\sum_{i=1}^{k} x_{i.} = \frac{1}{kn}\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}$$

The total sum of squares S is defined to be $\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2$. This sum of squares can be partitioned as follows:

$$S = \sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2 = knx_{..}^2 + n\sum_{i=1}^{k}(x_{i.} - x_{..})^2 + k\sum_{j=1}^{n}(x_{.j} - x_{..})^2$$
$$+ \sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - x_{i.} - x_{.j} + x_{..})^2 \qquad 1.2$$

where the four quantities on the right of equation 1.2 are respectively the sum of squares (M) assignable to the overall mean, blocks (A), treatments (B) and residual (C).

The usual analysis of variance ratio for testing the null hypothesis H_0: the effect assignable to treatment j is zero; versus the alternative hypothesis H_1: the effect assignable to treatment j is not zero; for all j is

$$\frac{k \sum_{j=1}^{n} (x_{.j} - x_{..})^2/(n-1)}{\sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - x_{i.} - x_{.j} - x_{..})^2/(n-1)(k-1)} \qquad 1.3$$

In order for this ratio to have an F distribution with $(n-1)$ and $(k-1)(n-1)$ degrees of freedom, it is sufficient that the x_{ij} be random independent observations from normal populations with the same variances. It is also necessary to assume that the effects of blocks and treatments are additive. Analytically, this last assumption states that the means of the normal population associated with the individual cells are assumed to be of the form:

$$u_{ij} = \mu + \beta_i + \tau_j$$

with

$$\Sigma\beta_i = 0 \quad \text{and} \quad \Sigma\tau_j = 0.$$

The parameter μ is the average of all the population means, β_i is the average effect assignable to block i. The assumption of additivity implies that if block i gives rise to a measurement ten units greater than the measurement on the same treatment in some other block, then any other measurement in block i will be ten units larger (in the population mean) than the corresponding measurement in the other block.

It can be shown that the sums of squares on the right hand side of equation 1.2, when divided by σ^2, the common population variance, are each independently distributed by chi-square laws with $(k-1)(n-1)$, $(k-1)$ and $(n-1)$ and 1 degrees of freedom respectively. It follows, then, that under suitable assumptions expression 1.3 is distributed according to an F with $(n-1)$ and $(k-1)(n-1)$ degrees of freedom.

Instead of the F ratio, Pitman [12] considered the ratio

$$W = \frac{\Sigma(\text{Treatments})^2}{\Sigma(\text{Treatments})^2 + \Sigma(\text{Residual})^2}$$

This ratio will now be denoted by $W = B/(B + C)$. Under the above assumptions B and C when divided by σ^2 are independently distributed as χ^2's with $(n-1)$ and $(k-1)(n-1)$ degrees of freedom respectively. It follows that W has a beta distribution with $(n-1)$ and $(k-1)(n-1)$ degrees of freedom.

Pitman [12] treated the problem of testing the null hypothesis that the treatments are all equal without making any assumption about the x_{ij}. If the null hypothesis is in fact true, the value of W observed is the result of the treatments, now mere labels, being distributed at random to the various individuals in the blocks.

An equally likely value for W would arise if the observed responses were shuffled within the blocks. Consider for example the case of three treatments and three blocks where the data are given in Table 1.1(a). The value of W from this set of observations is 21/65. If the null hypothesis is true then the observed values 4, 2, -3 in Block 1 might have appeared in the order 2, -3, 4, say, and Blocks

2 and 3 might have read 1, -2, 2 and 0, -1, 3 respectively as given in Table 1.1(b). After this shuffling, $W = 57/65$. It is to be noted that the block sums of squares remains unchanged under such reshuffling. The sums of squares which do vary are treatment and

Table 1.1

	Treatments				Treatments		
Blocks	1	2	3	Blocks	1	2	3
1	4	2	-3	1	2	-3	4
2	2	-2	1	2	1	-2	2
3	0	3	-1	3	0	-1	3
Treatment sum of squares:			14			38	
Residual sum of squares:			88/3			16/3	
	(a)				(b)		

residual. The number of different values that W can take on for this example is exactly the number of ways in which the blocks can be reshuffled in such a way that the treatment means are different from every other set of treatment means obtained from the other shufflings. The number of values of W, each value being equally likely, for the case of n treatments and k blocks is the number of ways that n numbers can be arranged $(n!)$ raised to the power k, the number of blocks. These $(n!)^k$ values of W contain multiplicities because many of the shufflings will give rise to what is equivalent to interchanging the columns of numbers in Table 1.1(a). The number of such different interchanges is $n!$. Thus the number of distinct equally likely values of W is at most $(n!)^k/n!$ or $(n!)^{k-1}$. At most refers to accidental equalities of W due to the data being rounded. An exact test for the null hypothesis would be to compare the actual observed value of W from the experiment with all the other values of W from the $(n!)^{k-1}$ shufflings. If not more than some pre-assigned $\alpha\%$ of these $(n!)^{k-1}$ values of W exceed the experimental value one rejects the null hypothesis at the $\alpha\%$ level. To avoid the computation required to find all these values of W, Pitman investigated the distribution of W when the underlying distribution deviates from normality. There are two different distributions of W being discussed. One is the distribution of W which is associated with the underlying distribution and hence does not depend on the individual values of the sample. The other is the distribution of W obtained from the permutations and depends directly on the sample. These distributions will be denoted by unconditional and conditional respectively.

Under the assumption that the underlying distribution is normal, the unconditional distribution of W is a beta with $(n-1)$ and $(k-1)(n-1)$ degrees of freedom. If W deviated very little from the beta distribution for any reasonable underlying distribution, then one could "assume normality" and proceed to use the F test without too much concern, since a "$\alpha\%$ test" would, in general, reject really true null hypotheses approximately $\alpha\%$ of the time.

It was shown by Eden and Yates [3] that there was good agreement between the beta distribution and the conditional distribution of W in a particular sampling experiment. Pitman [12] obtained the first four permutation moments of W in terms of the first four moments and cumulants of the observations. Welch [14], in the paper that parallels Pitman's, obtained only the first two moments. By looking at the moments, Pitman was able to draw certain conclusions about the distribution of W as the underlying distribution deviates from normality. Pitman's moments can be made to look simpler and hence easier to study by introducing a change of notation. Let

$$\sigma_j^2 = \frac{\sum_{i=1}^{k}(x_{ij} - x_{.j})^2}{k-1}$$

$$k_{3j} = g_j\sigma_j^3 \qquad \text{where} \qquad k_{3j} = \frac{\sum_{i=1}^{k}(x_{ij} - x_{.j})^3}{k-1}$$

$$k_{4j} = G_j\sigma_j^4 \qquad \text{where} \qquad k_{4j} = \frac{\sum_{i=1}^{k}(x_{ij} - x_{.j})^4}{k-1} - 3\sigma_j^2$$

The notation Σ'' is used to denote summation over all subscripts contained inside the summation omitting all terms of the summation for which the subscripts are equal. Further "avep" is used to indicate that the average is being taken over permutations within blocks. With these notational changes Pitman's moments now appear as

$$\operatorname{avep}(W) = \frac{1}{k}$$

$$\operatorname{avep}((W - \operatorname{avep}(W))^2) = \frac{2}{k^2(n-1)} \frac{\sum_1^{k}{}'' \sigma_p^2\sigma_q^2}{(\sum_1^k \sigma_p^2)^2}$$

$$\operatorname{avep}((W - \operatorname{avep}(W))^3) = \frac{8}{k^3(n-1)^2} \frac{\sum_1^{k}{}'' \sigma_p^2\sigma_q^2\sigma_r^2}{(\sum_1^k \sigma_p^2)^3} + \frac{4(n-1)(n-2)}{k^3n^4} \frac{\sum_1^{k}{}'' g_p g_q \sigma_p^3\sigma_q^3}{(\sum_1^k \sigma_p^2)^3}$$

$$\text{avep}((W - \text{avep}(W))^4) = \frac{12}{k^4(n-1)^2}\,\frac{(\sum_1^{k}{}'' \sigma_p^2\sigma_q^2)^2}{(\sum_1^k \sigma_p^2)^4} - \frac{48}{k^4(n-1)^2(n+1)}\,\frac{\sum_1^{k}{}'' \sigma_p^2\sigma_q^2}{(\sum_1^k \sigma_p^2)^4}$$

$$+ \frac{48}{k^4(n-1)^3}\,\frac{\sum_1^{k}{}'' \sigma_p^2\sigma_q^2\sigma_r^2\sigma_s^2}{(\sum_1^k \sigma^2)^4} + \frac{8(n-1)(n-2)(n-3)}{k^4 n^5(n+1)}\,\frac{\sum_1^{k}{}'' G_p G_q \sigma_p^4\sigma_q^4}{(\sum_1^k \sigma_p^2)^4}$$

$$+ \frac{96(n-2)}{k^4 n^4}\,\frac{\sum_1^{k}{}'' \varepsilon_p \varepsilon_q \sigma_p^3\sigma_q^3 \sum_1^k \sigma_r^2}{(\sum_1^k \sigma_p^2)^4}$$

These formulas can be simplified somewhat by letting $\tau_i = \dfrac{\sigma_i}{\sqrt{\sum_1^k \sigma_p^2}}$.

Also instead of giving the first four moments of W we will write down cumulants of kW. These are the average, variance, skewness and elongation (or kurtosis).

$$\text{avep}(kW) = 1$$

$$\text{varp}(kW) = \text{avep}((kW - \text{avep}(kW))^2) = \frac{1}{n-1}\sum_1^{k}{}'' \tau_p^2\tau_q^2$$

$$\text{skep}(kW) = \text{avep}((kW - \text{avep}(kW))^3) = \frac{8}{(n-1)^2}\sum_1^{k}{}'' \tau_p^2\tau_q^2\tau_r^2 + \frac{4(n-1)(n-2)}{n^4}\sum_1^{k}{}'' \varepsilon_p\varepsilon_q\tau_p^3\tau_q^3$$

$$\text{elop}(kW) = \text{avep}((kW - \text{avep}(kW))^4) - 3(\text{avep}((kW - \text{avep}(kW))^2))^2$$

$$= \frac{-48}{(n-1)^2(n+1)}\sum_1^{k}{}'' \tau_p^2\tau_q^2 + \frac{48}{(n-1)^3}\sum_1^{k}{}'' \tau_p^2\tau_q^2\tau_r^2\tau_s^2 +$$

$$\frac{8(n-1)(n-2)(n-3)}{n^5(n+1)}\sum_1^{k}{}'' G_p G_q \tau_p^4\tau_q^4 + \frac{96(n-2)}{n^4}\sum_1^{k}{}'' \varepsilon_p\varepsilon_q\tau_p^3\tau_q^3 \sum_1^k \tau_r^2$$

Of these four cumulants of W -- average, variance, skewness and elongation -- only the average is independent of the particular observations from the experiment.

The mean and variance of a beta distribution with $(n-1)$ and $(k-1)(n-1)$ degrees of freedom are $\frac{1}{k}$ and $\frac{2(k-1)}{k^2(kn - k+2)}$ respectively. The average value of W obtained from randomization is thus the same as the average value for the beta distribution with the proper number of degrees of freedom. When the variance of the beta distribution is equated to the variance of W one gets the condition

$$\frac{2}{n-1} \sum_{1}^{k}{}'' \tau_p^2 \tau_q^2 = \frac{2(k-1)}{kn - k+2}$$

When all the block variances are equal then $\tau_p^2 = \tau_q^2 = \frac{1}{k}$ for $1 \leq p, q \leq k$. Thus the left hand side for equal block variances becomes

$$\frac{2\, k(k-1)}{(n-1)\, k^2} = \frac{2(k-1)}{k(n-1)}.$$

The right hand side of the equation can be written as $\frac{2(k-1)}{k(n-1) + 2}$ which is always smaller than $\frac{2(k-1)}{k(n-1)}$ but by an amount which becomes small as k and n increase. The value of the left hand side decreases from $\frac{2(k-1)}{k(n-1)}$ to zero as one block variance becomes much larger than the others. Thus the variance of W is not greater than $\frac{2(k-1)}{k^3(n-1)}$ and it takes this value when the block variances are equal. When the variance of W is near a maximum, the conditional variance agrees well with the variance of the appropriate beta distribution.

If a few of the block variances are very large relative to the others, then the conditional variance of W will be smaller than the variance of the corresponding beta distribution. Pitman suggests three possibilities when this arises:

(i) Discard the blocks which have large variance relative to the other blocks.

(ii) Fit a beta distribution by the use of the first two moments of W (investigated by Welch [14]). This will mean a decrease in the number of degrees of freedom in the beta distribution which applies under normality.

(iii) Make all block variances equal. This of course requires the calculation of each of the block variances and the adjustment of the observations in each block to make the necessary changes in the block variances.

One should not consider doing (i), (ii), or (iii) if suitable normality assumptions can be made, even if the block variances seem to be very unequal.

Shewhart and Winters [9] tested Student's z distribution for the two-sample problem by taking samples from normal, rectangular and right-triangular distributions and found that there was only satisfactory agreement for the case of the sample from the normal population. Pearson [7] did a similar analysis for samples drawn from several non-normal Pearsonian distributions but his results were somewhat inconclusive. Rider [8] has treated the case of samples from U-shaped distributions and Baker [1] has similarly treated theoretical non-homogeneous populations.

If the block variances σ_i^2 are all equal, or approximately equal, the second moment of W under randomization theory will be too large, but this has an appreciable effect only when k and n are small. The variance of W is fairly insensitive to changes in the values of the block variances when k is large. Pitman recommends that the second moment of W be calculated and compared with the variance of the appropriate beta distribution when k or n is less than 5.

Pitman further remarks that if a beta distribution is fitted by means of the first two moments of W, the third and fourth moments are likely to agree well provided that the second moment is ot too small. For equal block variances, fitting a beta distribution in this manner gives a good approximation provided $k(n-1)$ is not too small.

3. Some Further Remarks

Besides these results of Pitman, other writers have investigated the analysis of variance procedure when the underlying distribution is non-normal.

It is often suggested that before making an analysis of variance test it is wise first to make a test for homogeneity of variances (e.g. Bartlett's test). Box [2] shows that when little is known of the form of the underlying distribution, making such a test is often more dangerous than omitting it. Box further states that when there are unequal treatment sizes k_i and the variances from treatment to treatment are suspected of being unequal, then it seems that the usual analysis of variance procedure that says to pool the within treatment estimates should be replaced by a criterion that employs a weighted estimate of between treatment variance as suggested by Welch [15] and by James [6]. To illustrate this, assume a situation in which we have n treatments, the i^{th} treatment having k_i units. If s_i^2 is the estimate of variance for the i^{th} treatment, $x_{i.}$ is the mean of the i^{th} treatment and $x_{..} = \sum_{i=1}^{n} \sum_{j=1}^{k} x_{ij} / \sum_{i=1}^{k_i} k_i$ is the overall mean, then the criterion would be $\sum_{i=1}^{n} v_i(x_{i.} - x_{..})^2$ where $v_i = k_i/s_i^2$, in place of the usual pooled estimate criterion $\sum_{i=1}^{n}(x_{i.} - x_{..})^2/s^2$ where $s^2 = \sum_{i=1}^{n}(k_i - 1)s_i^2 / \sum_{i=1}^{n} k_i$. In fact Box suggests that the criterion based on a weighted estimate of the between treatment variance might even be used for the case of equal groups where heterogeneity of variance might occur. The criterion would be $\sum_{i=1}^{n}(x_{i.} - x_{..})^2/s_i^2$ instead of the usual $\sum_{i=1}^{n}(x_{i.} - x_{..})^2/s^2$ because it seems more reasonable to weigh those treatment means which have large within treatment variances less than those treatments which have small within treatment variances. This criterion seems sensible to use if one first computed the covariance between $(x_{i.} - x_{..})^2$ and s_i^2 and found it to be positive, as, it is hoped, would usually be the case.

Chapter 2

DYADS AND DYADICS FOR THE MULTIDIMENSIONAL PROBLEM

2.1 Two Treatment Multidimensional Case

Consider a two-way layout as in Chapter 1 with k blocks but with only two treatments per block. Suppose now that several characteristics are measured on each individual, instead of but one, as was the case with the work of Pitman and Welch. The two-way classification can be displayed as follows:

	TREATMENTS	
BLOCKS	1	2
1	x_{11}	x_{12}
2	x_{21}	x_{22}
⋮	⋮	⋮
k	x_{k1}	x_{k2}

where x_{ij}; $i = 1, 2, \ldots, k$; $j = 1, 2$ is the measurement of the response of the j^{th} treatment applied to an element in the i^{th} block and is, in general, a m-dimensional vector. The k^{th} component of x_{ij} is the response of the k^{th} character measured on the i^{th} block subject to the j^{th} treatment. Detailed consideration here is limited to only two treatments, but the methods used should generalize to include more than two treatments. Thus it might be possible to find results for multidimensional analysis of variance that are suitably analogous to Pitman's results in the univariate case.

2.2 Multidimensional Mathematical Tools

One problem that immediately arises in the multidimensional case is that of understanding the meaning of what, in a single dimension, was a sum of squares. For example, we have to be able to generalize the meaning of such things as $\sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - x_{..})^2$ when the x_{ij} and $x_{..}$ are m-dimensional vectors. It is also necessary to define them in such a way that they are convenient to use. One way of defining sums of squares of vectors was treated by Tukey [13]. He makes use of quantities called (by Gibbs [4]) dyads and dyadics. Dyads are formed by the "product" of vectors in a way best shown by the following example:

$$(a_1, b_1)*(a_2, b_2) = \begin{pmatrix} a_1a_2 & a_1b_2 \\ a_2b_1 & b_1b_2 \end{pmatrix}$$

The sum of two or more dyads is called a dyadic. This addition to form a dyadic is performed componentwise:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ d_1 & e_1 & f_1 \\ g_1 & h_1 & i_1 \end{pmatrix} + \begin{pmatrix} a_2 & b_2 & c_2 \\ d_2 & e_2 & f_2 \\ g_2 & h_2 & i_2 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 & c_1 + c_2 \\ d_1 + d_2 & e_1 + e_2 & f_1 + f_2 \\ g_o + g_2 & h_1 + h_2 & i_1 + i_2 \end{pmatrix}$$

These operations need not of course be restricted to two or three dimensional vectors. For an analysis of variance with vectors instead of scalars, we may replace the sum of squares by a dyadic of sums of squares and cross-products. The sums of squares resulting from a set of k m-dimensional vectors will be a dyadic with m^2 components.

As an example consider a design consisting of only two experiments with two measurements being taken in each experiment. Suppose that the results are as follows:

Character	Experiment I	Experiment II
A	5	3
B	3	1

The dyad which is the contribution to the total sums of squares and cross products from the first experiment is the product of the vector (5, 3) with itself. Thus the total sums-of-squares-and-cross-products S is the dyadic with four components

$$S = (5,3).(5,3) + (2,1).(2,1) = \begin{pmatrix} 25 & 15 \\ 15 & 9 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 29 & 17 \\ 17 & 10 \end{pmatrix}$$

2.3 Dyads as Applied to a 2 x k Experiment in m Dimensions

We will now proceed to apply these rules of operation of dyads and dyadics to the $2k$ vectors given in section 2.1. Let c_i be the difference vector $(x_{i1} - x_{i2})$. Furthermore let $\epsilon_i c_i$ for $1 \leq i \leq k$ be the set of difference vectors when the x_{ij} are permuted within block i. In this way, $\epsilon_i = +1$ if the difference vector is as found from the observed data and -1 if the difference vector is as found after interchanging the vectors in block i. The total "sum of squares" of the differences for such an arrangement can be partitioned as follows:

$$\sum_{i=1}^{k}(\epsilon_i c_i)^2 = \frac{1}{k}\left(\sum_{i=1}^{k}\epsilon_i c_i\right)^2 + \text{RESIDUAL}$$

Let $A = \sum_{i=1}^{k}(\epsilon_i c_i)^2$, $B = \frac{1}{k}\left(\sum_{i=1}^{k}\epsilon_i c_i\right)^2$, and C = RESIDUAL

Further let us denote the components of the m-dimensional vector $\epsilon_i c_i$ by $(\epsilon_i a_{i1}, \epsilon_i a_{i2}, \ldots, \epsilon_i a_{im})$. Then A is a dyadic with m^2 terms obtained by adding k dyads, the i^{th} one of which is

$$\begin{pmatrix} a_{i1} & a_{i1}a_{i2} & \cdots & a_{i1}a_{im} \\ a_{i1}a_{i2} & a_{i2}^2 & \cdots & a_{i2}a_{im} \\ \vdots & \vdots & & \vdots \\ a_{i1}a_{im} & a_{i2}a_{im} & & a_{im}^2 \end{pmatrix}$$

Thus

$$A = \begin{pmatrix} \sum_{i=1}^{k} a_{i1}^2 & \sum_{i=1}^{k} a_{i1}a_{i2} & \cdots & \sum_{i=1}^{k} a_{i1}a_{im} \\ \sum_{i=1}^{k} a_{i1}a_{i2} & \sum_{i=1}^{k} a_{i2}^2 & \cdots & \sum_{i=1}^{k} a_{i2}a_{im} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^{k} a_{i1}a_{im} & \sum_{i=1}^{k} a_{i2}a_{im} & \cdots & \sum_{i=1}^{k} a_{im}^2 \end{pmatrix}$$

B, on the other hand, is a dyad which is formed by taking $\frac{1}{k}$ times the product of a vector D by itself where D is obtained by summing the m-dimensional vectors $\epsilon_i c_i$ over the k blocks.

$$D = \left(\sum_{i=1}^{k} \epsilon_i a_{i1}, \ \sum_{i=1}^{k} \epsilon_i a_{i2}, \ \ldots, \ \sum_{i=1}^{k} \epsilon_i a_{im}\right)$$

$$B = \frac{1}{k} D.D = \begin{pmatrix} \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{i1}\right)^2}{k} & \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{i1}\right)\left(\sum_{i=1}^{k} \epsilon_i a_{i2}\right)}{k} & \cdots & \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{i1}\right)\left(\sum_{i=1}^{k} \epsilon_i a_{im}\right)}{k} \\ \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{i1}\right)\left(\sum_{i=1}^{k} \epsilon_i a_{i2}\right)}{k} & \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{i2}\right)^2}{k} & \cdots & \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{i2}\right)\left(\sum_{i=1}^{k} \epsilon_i a_{im}\right)}{k} \\ \vdots & \vdots & & \vdots \\ \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{i1}\right)\left(\sum_{i=1}^{k} \epsilon_i a_{im}\right)}{k} & \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{i2}\right)\left(\sum_{i=1}^{k} \epsilon_i a_{im}\right)}{k} & \cdots & \frac{\left(\sum_{i=1}^{k} \epsilon_i a_{im}\right)^2}{k} \end{pmatrix}$$

Pitman considered, in the univariate case, what would, in the present notation, be denoted by $B/(B + C)$ where each $\epsilon_i c_i$ would then be a scalar. Since $B + C = A$, the ratio may also be written B/A. This ratio in the univariate case is merely a (dimensionless) numerical quantity. The "ratio" which will be used here in the multi-dimensional situation is B/A looked at in such a way that A remains constant. A convenient constant for A is the $m \times m$ dyadic whose diagonal elements are all one and whose off diagonal elements are all zero. A further discussion of this "ratio" B/A is given in Chapter 3. It would then be enough to compute moments of this new dyad. Furthermore, one would not need to attempt to attach a meaning to the "ratio" of an arbitrary dyad and an arbitrary dyadic.

This simplification requires that $\epsilon_i a_{ij}$, $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, m$ be replaced by new numbers $\epsilon_i a'_{ij}$ such that $\sum_{i=1}^{k} a'^2_{ip} = 1$ for $1 \leq p \leq m$ and $\sum_{i=1}^{k} a'_{ip} a'_{iq} = 0$ for $1 \leq p \neq q \leq m$. The class of transformations which meet these requirements is a linear one. For the special case of two dimensions a transformation satisfying these requirements is given by

$$a'_{11} = \frac{a_{11}}{\sqrt{\sum_{p=1}^{k} a_{p1}^2}} \quad \text{and} \quad a'_{12} = \frac{a_{12} - \dfrac{\sum_{p=1}^{k} a_{p1}a_{p2}}{\sum_{p=1}^{k} a_{p1}^2}\, a_{11}}{\sqrt{\sum_{q=1}^{k}\left(a_{q2} - \dfrac{\sum_{p=1}^{k} a_{p1}a_{p2}}{\sum_{p=1}^{k} a_{p1}^2}\, a_{q1}\right)^2}} \qquad 2.1$$

This is only one of many transformations such that $\sum_{p=1}^{k} a'^{2}_{p1} = \sum_{p=1}^{k} a'^{2}_{p2} = 1$ and $\sum_{p=1}^{k} a'_{p1}a'_{p2} = 0$. The class of transformations, of which the transformation given by equations 2.1 is a member, consists of all a^*_{p1}, a^*_{p2} such that

$$\begin{aligned} a^*_{p1} &= a'_{p1} \cos\theta - a'_{p2} \sin\theta \\ a^*_{p2} &= a'_{p2} \sin\theta + a'_{p2} \cos\theta \end{aligned} \qquad 2.2$$

where $0 \leq \theta \leq 2\pi$. Mathematical definitions for the class of transformations for the general m-dimensional situation follow directly from the two dimensional definition given by 2.1 and 2.2.

Chapter 3

DISCUSSION OF THE CRITERION B/A

3.1 Partitioning of the Sum of Squares

Consider for the present the general univariate case of n treatments in k blocks as discussed by Pitman. We are given the measurements

$$\begin{matrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{matrix}$$

where x_{ij} is the observed response of the i^{th} block subjected to the j^{th} treatment. The most general quadratic form relating these x_{ij}'s is given by

$$Q = \sum_{i=1}^{k} \sum_{j=1}^{n} \sum_{I=1}^{k} \sum_{J=1}^{n} a_{ijIJ} x_{ij} x_{IJ} \qquad 3.1$$

In order for Q to be invariant under interchanges (shuffling) of blocks among themselves and of treatments among themselves, certain conditions involving the a_{ijIJ} are necessary. To determine these conditions, suppose that jJ is fixed. All terms in Q for which this is true are

$$\sum_{i=1}^{k} \sum_{I=1}^{k} a_{ijIJ} x_{ij} x_{IJ} = \sum_{i=1}^{k} a_{ijiJ} x_{ij} x_{iJ} + \sum_{i,I=1}^{k}{}'' a_{ijIJ} x_{ij} x_{IJ} \qquad 3.2$$

Suppose all but two of the x_{ij} are zero. Further, suppose that the two non-zero x_{ij}'s appear in the same block. Let us denote these

x_{ij} by x_{pj} and x_{pJ}. The first summation (let it be denoted by Σ_1) on the right hand side of equation 3.2 will have exactly one non-zero term $a_{pjpJ}x_{pj}x_{pJ}$. Suppose now the blocks are shuffled so that block p no longer contains the two non-zero x_{ij}'s. The value of the only non-zero term in Σ_1 will now be $a_{p'jp'J}x_{p'j}x_{p'J}$ where block p' is now the block containing the non-zero x_{ij}'s. Unless $a_{pjpJ} = a_{p'jp'J} = a_{jJ}$, Q will not remain unchanged under shuffling of the blocks. If we now suppose again that only two of the x_{ij}'s are non-zero but this time assume that these two non-zero x_{ij}'s are in two separate blocks, the second summation (denoted by Σ_2) on the right hand side of equation 3.2 will have exactly two non-zero terms $a_{pjqJ}x_{pj}x_{qJ}$ and $a_{qJpJ}x_{qJ}x_{pJ}$. Again, supposing that the blocks are shuffled, it is necessary that the coefficients a_{pjqJ} and a_{qjpJ} be constant for fixed jJ. Thus it is necessary that equation 3.2 be of the form

$$\sum_{i=1}^{k}\sum_{I=1}^{k} a_{ijIJ}x_{ij}x_{IJ} = a_{jJ}\sum_{i=1}^{k} x_{ijiJ} + b_{jJ}\sum_{i,I=1}^{k}{}'' x_{ij}x_{IJ} \qquad 3.3$$

By reversing the roles of the blocks and treatments and repeating the above argument it follows that a general quadratic form satisfying the required symmetry with respect to blocks and treatments can be partitioned as follows:

$$Q = a_1\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2 + b_1\sum_{j=1}^{n}\sum_{i,I=1}^{k}{}'' x_{ij}x_{Ij} + c_1\sum_{i=1}^{k}\sum_{j,J=1}^{n}{}'' x_{ij}x_{iJ} + d_1\sum_{i,I=1}^{k}{}''\sum_{j,J=1}^{n}{}'' x_{ij}x_{IJ} \qquad 3.4$$

3.2 The Adequacy of Studying B/A, provided A is scaled as 1.

From the preceding arguments on the partitioning of a quadratic form it follows that the total sum of squares of the x_{ij}'s can be partitioned into four parts. With some algebraic manipulation the partitioning can be reduced to a sum of squares involving the mean of the x_{ij}'s, one involving the means of the treatments $(x_{.j})$, one involving the means of the blocks $(x_{i.})$ and finally one involving both the treatment means and the block means (residual or error sum of squares). Without any prior information on the mean of the underlying population, any criterion which tests for the difference between treatments should be independent of the overall mean. Since, moreover, randomizations within blocks leave the block sum of squares unchanged, we need concern ourselves only with a criterion involving the treatment sum of squares and the error sum of squares. These two sums of squares were denoted by B and C in Chapter 2. A ratio involving B and C is chosen arbitrarily, following the Pitman approach. The most general ratio involving B and C is of the form $(\alpha B + \beta C)/(\gamma B + \delta C)$.

Consider for the purpose of illustration the special case where only two treatments are under test and where only a single response is being measured. Suppose that we scale the measurements such that $B + C = 1$. We will denote B by B_1 and C by C_1 under this scaling.

Suppose an experiment consists of five blocks with two treatments per block and one experiment on each treatment in each block. The responses can be expressed by five numbers which are the differences of the yields in the five blocks. Furthermore, these five differences, call them a_i, can be displayed as points on a line as shown in Figure 3.1. If we now change all the signs of the a_i, we obtain five more

Figure 3.1

points to place on the line which are the same distance from 0 in the opposite direction as shown by Figure 3.2. By a rescaling of

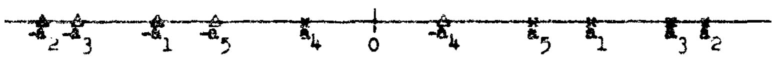

Figure 3.2

these differences a_i we can find a unit of measurement such that

$$\sum_{i=1}^{5} a_i^2 = 1.$$

When this has been done, any particular randomization R giving rise to differences $(\epsilon_1 a_1, \epsilon_2 a_2, \epsilon_3 a_3, \epsilon_4 a_4, \epsilon_5 a_5)$, where ϵ_i is +1 or -1 (depending on whether a_i occurs without or with a change of sign) is such that $\sum_{i=1}^{5}(\epsilon_i a_i)^2 = 1$. The sum of squares $\sum_{i=1}^{5}(\epsilon_i a_i)^2$ is B + C and any criterion based on the a_i other than one which has B + C = 1 (or some other convenient constant) or a multiple of B + C in the denominator will have a different denominator for each randomization chosen. The numerator can have any linear combination of 1 and $\frac{1}{5}(\sum_{i=1}^{5}\epsilon_i a_i)^2 = \beta$. A reasonable criterion is thus of the form $(\alpha B + \beta C)/(B + C)$. However, by finding the randomization cumulants for a special case of this criterion, namely $B/(B + C)$ or B/A scaled such that $A = 1$, one can then find the randomization cumulants for the more general criterion $(\alpha B + \beta C)/A$. The formulas relating the randomization cumulants for the general criterion $(\alpha B + \beta C)/A$ in terms of the cumulants for B/A are now derived.

$$\begin{aligned}
\text{avep}\{\alpha B + \beta C\} &= \alpha\ \text{avep}\{B\} + \beta\ \text{avep}\{C\} \\
&= \alpha\ \text{avep}\{B\} + \beta\ \text{avep}\{1 - B\} \\
&= (\alpha - \beta)\ \text{avep}\{B\} + \beta \\
\text{varp}\{\alpha B + \beta C\} &= \text{varp}\{(\alpha - \beta)B\} \\
&= (\alpha - \beta)^2\ \text{varp}\{B\} \\
\text{skep}\{\alpha B + \beta C\} &= \text{skep}\{(\alpha - \beta)B\} \\
&= (\alpha - \beta)^3\ \text{skep}\{B\} \\
\text{elop}\{\alpha B + \beta C\} &= \text{elop}\{(\alpha - \beta)B\} \\
&= (\alpha - \beta)^4\ \text{elop}\{B\}
\end{aligned}$$

3.3 Multivariate Analog of B/A

In the multiresponse situation the condition $B + C = 1$ becomes $B + C = I$ where, for the case of m responses, I is a $m \times m$ dyad with 1's on the diagonal and 0's elsewhere. Again it will make no difference whether we are considering B_I (B under the condition $B + C = I$) or C_I or some linear combination of B_I and C_I. As was the case in the univariate situation, the randomization cumulants for the criterion B with suitable scaling so that $A = I$ are sufficient for obtaining the cumulants of any arbitrary linear combination of B and C.

$$\begin{aligned}
\text{avep}(\alpha B + \beta C) &= \text{avep}(\alpha B + \beta(I - B)) \\
&= \beta I + \text{avep}((\alpha - \beta)B) \\
&= \beta I + (\alpha - \beta)\ \text{avep}(B)
\end{aligned}$$

$$\begin{aligned}
\text{varp}(\alpha B + \beta C) &= \text{varp}(\alpha B + \beta(I - B)) \\
&= \text{varp}((\alpha - \beta)B + \beta I) \\
&= \text{varp}((\alpha - \beta)B) \\
&= \text{avep}((\alpha - \beta)B.(\alpha - \beta)B) - \text{avep}((\alpha - \beta)B).\text{avep}((\alpha - \beta)B) \\
&= \text{avep}((\alpha - \beta)^2 B.B) - (\alpha - \beta)^2\ \text{avep}(B).\text{avep}(B) \\
&= (\alpha - \beta)^2[\text{avep}(B.B) - \text{avep}(B).\text{avep}(B)] \\
&= (\alpha - \beta)^2\ \text{varp}(B)
\end{aligned}$$

$$\begin{aligned}
\text{skep}(\alpha B + \beta C) &= \text{skep}(\alpha B + \beta(I - B)) \\
&= \text{skep}((\alpha - \beta)B + \beta I) \\
&= \text{skep}((\alpha - \beta)B) \\
&= \text{avep}([(\alpha - \beta)B - \text{avep}((\alpha - \beta)B)].[(\alpha - \beta)B - \text{avep}((\alpha - \beta)B)]. \\
&\qquad [(\alpha - \beta)B - \text{avep}((\alpha - \beta)B)]) \\
&= \text{avep}((\alpha - \beta)[B - \text{avep}(B)].(\alpha - \beta)[B - \text{avep}(B)]. \\
&\qquad (\alpha - \beta)[B - \text{avep}(B)]) \\
&= (\alpha - \beta)^3\text{avep}((B - \text{avep}(B)).(B - \text{avep}(B)).(B - \text{avep}(B))) \\
&= (\alpha - \beta)^3\ \text{skep}(B)
\end{aligned}$$

Finally

$$\begin{aligned}
\text{elop}(\alpha B + \beta C) &= \text{elop}(\alpha B + \beta(I - B)) \\
&= \text{elop}((\alpha - \beta)B + \beta I) \\
&= \text{elop}((\alpha - \beta)B) \\
&= \text{avep}([(\alpha-\beta)B - \text{avep}((\alpha-\beta)B)].[(\alpha-\beta)B - \text{avep}((\alpha-\beta)B)]. \\
&\qquad [(\alpha-\beta)B - \text{avep}((\alpha-\beta)B)].[(\alpha-\beta)B - \text{avep}((\alpha-\beta)B)]) \\
&\qquad - 3\ \text{varp}((\alpha-\beta)B).\text{varp}((\alpha-\beta)B)
\end{aligned}$$

$$= \operatorname{avep}([(\alpha-\beta)[B - \operatorname{avep}(B)]\cdot(\alpha-\beta)[B - \operatorname{avep}(B)]\cdot(\alpha-\beta)[B - \operatorname{avep}(B)]\cdot$$
$$(\alpha-\beta)[B - \operatorname{avep}(B)]]) - 3(\alpha-\beta)^2 \operatorname{varp}(B)\cdot(\alpha-\beta)^2 \operatorname{varp}(B)$$
$$= (\alpha-\beta)^4 \operatorname{avep}[(B - \operatorname{avep}(B))\cdot(B - \operatorname{avep}(B))\cdot(B - \operatorname{avep}(B))\cdot$$
$$(B - \operatorname{avep}(B))] - 3(\alpha-\beta)^4 \operatorname{varp}(B)\cdot\operatorname{varp}(B)$$
$$= (\alpha - \beta)^4 \operatorname{elop}(B).$$

The derivation of the first four cumulants of B for all permutations of sign is given in Chapter 4. In keeping with the notation of Chapter 2, the dyad B is formed by taking the product of the vector of sums of differences, $\sum_{i=1}^{k} c_i$, with itself, divided by k, the number of blocks. This can be written as

$$B = \frac{1}{k}\left(\sum_{i=1}^{k} c_i\right)\left(\sum_{i=1}^{k} c_i\right)$$

The results of Chapter 4 are:

$$\operatorname{avep}(B) = \frac{1}{k}\sum_{p=1}^{k} c_p\cdot c_p = \frac{1}{k} A$$

$$\operatorname{varp}(B) = \frac{1}{k^2}\sum_{p,q=1}^{k}{}'' (c_p\cdot c_q\cdot c_p\cdot c_q + c_p\cdot c_q\cdot c_q\cdot c_p)$$

$$\operatorname{skep}(B) = \frac{1}{k^3}\sum_{p,q,r=1}^{k}{}'' (c_p\cdot c_q+c_q\cdot c_p)\cdot(c_q\cdot c_r+c_r\cdot c_q)\cdot(c_r\cdot c_p+c_p\cdot c_r)$$

$$\operatorname{elop}(B) = \frac{1}{k^4}\{\sum_{p,q,r,s=1}^{k}{}'' [(c_p\cdot c_q+ c_q\cdot c_p)\cdot(c_q\cdot c_r+ c_r\cdot c_q)\cdot(c_r\cdot c_s+ c_s\cdot c_r)\cdot$$
$$(c_s\cdot c_p+ c_p\cdot c_s) + (c_p\cdot c_q+ c_q\cdot c_p)\cdot(c_s\cdot c_p+ c_p\cdot c_s)\cdot(c_q\cdot c_r+ c_r\cdot c_q)\cdot$$
$$(c_r\cdot c_s+ c_s\cdot c_r) + (c_p\cdot c_q+ c_q\cdot c_p)\cdot(c_r\cdot c_s+ c_s\cdot c_r)\cdot(c_s\cdot c_p+ c_p\cdot c_s)\cdot$$
$$(c_q\cdot c_r+ c_r\cdot c_q)] - \sum_{p,q=1}^{k}{}''(c_p\cdot c_q+ c_q\cdot c_p)\cdot(c_p\cdot c_q+ c_q\cdot c_p)\cdot$$
$$(c_p\cdot c_q+ c_q\cdot c_p)\cdot(c_p\cdot c_q+ c_q\cdot c_p)\}.$$

3.4 Invariants and Cumulants over Rotation of B

In order to provide a procedure for testing whether, in any given instance, the two treatments are really alike, it is desirable to be able to compute a suitable quantity from the data and compare it with some predetermined value. The quantity that is computed in each instance should be independent of the particular choice of coordinates in terms of which the responses are expressed. The elements of the dyad B, even with the restriction that A be equal to a given constant, will depend on the choice of the transformation used to make A the particular constant we wish (most conveniently I). However the trace of B is independent of the choice of transformation which makes A equal to some convenient constant. This is now shown.

For convenience in notation let the observations after scaling to make $A = I$ be given by (a_i, b_i) in block i. Then

$$B = \frac{1}{k}\begin{pmatrix} (\sum_{i=1}^{k} a_i)^2 & (\sum_{i=1}^{k} a_i)(\sum_{i=1}^{k} b_i) \\ (\sum_{i=1}^{k} a_i)(\sum_{i=1}^{k} b_i) & (\sum_{i=1}^{k} b_i)^2 \end{pmatrix}$$

and

$$\text{trace } B = \frac{1}{k}\{(\sum_{i=1}^{k} a_i)^2 + (\sum_{i=1}^{k} b_i)^2\}$$

Let us now write trace B in a coordinate system (a', b') which has been rotated through an angle θ from (a, b). Then

$$a = a' \cos\theta - b' \sin\theta$$

$$b = a' \sin\theta + b' \cos\theta$$

Hence

$$\text{trace } B = [\sum_{i=1}^{k}(a_i' \cos\theta - b_i' \sin\theta)]^2 + [\sum_{i=1}^{k}(a_i' \sin\theta + b_i' \cos\theta)]^2$$

$$= (\sum_{i=1}^{k} a_i')^2 \cos^2\theta + (\sum_{i=1}^{k} b_i')^2 \sin^2\theta - 2(\sum_{i=1}^{k} a_i')(\sum_{i=1}^{k} b_i')\cos\theta \sin\theta$$

$$+ (\sum_{i=1}^{k} a_i')^2 \sin^2\theta + (\sum_{i=1}^{k} b_i')^2 \cos^2\theta + 2(\sum_{i=1}^{k} a_i')(\sum_{i=1}^{k} b_i')\cos\theta \sin\theta$$

$$= (\sum_{i=1}^{k} a_i')^2 + (\sum_{i=1}^{k} b_i')^2 .$$

Thus trace B does not depend on the angle θ through which the axes were rotated and is therefore independent of the orientation of the response variables.

To obtain the cumulants of trace B over permutations we need to define some new quantities. Let $L_\alpha B_\alpha$ denote a linear form of the elements of B. Here α is a double subscript. In particular, if L_α is such that $L_\alpha B_\alpha$ is trace B, then one only needs to express the permutation cumulants of $L_\alpha B_\alpha$ in terms of the permutation cumulants of B. Denoting the elements of L_α by L_{11}, L_{12}, L_{21}, L_{22} and those of B_α by B_{11}, B_{12}, B_{21}, B_{22}, we have

$$\text{avep}(L_\alpha B_\alpha) = \text{avep}(L_{11}B_{11} + L_{12}B_{12} + L_{21}B_{21} + L_{22}B_{22})$$

$$L_{11}\ \text{avep}(B_{11}) + L_{12}\ \text{avep}(B_{12}) + L_{21}\ \text{avep}(B_{21}) + L_{22}\ \text{avep}(B_{22}) \qquad 3.5$$

The right-hand side of 3.5 is equivalent to L_α operating on avep(B). Thus we have

$$\text{avep}(L_\alpha B_\alpha) = L_\alpha(\text{avep}(B))_\alpha . \qquad 3.6$$

Similarly $\quad \mathrm{varp}(L_\alpha B_\alpha) = L_\alpha L_\beta (\mathrm{varp}(B))_{\alpha\beta}$ 3.7

$$\mathrm{skep}(L_\alpha B_\alpha) = L_\alpha L_\beta L_\gamma (\mathrm{skep}(B))_{\alpha\beta\gamma} \qquad 3.8$$

$$\mathrm{elop}(L_\alpha L_\alpha) = L_\alpha L_\beta L_\gamma L_\delta (\mathrm{elop}(B))_{\alpha\beta\gamma\delta} \qquad 3.9$$

Thus the first four permutation cumulants of trace B can be obtained from 3.6, 3.7, 3.8 and 3.9 by replacing the operators L_α, L_β, L_γ and L_δ by an operator which has elements $L_{11} = L_{22} = 1$ and $L_{12} = L_{21} = 0$.

$$\mathrm{avep}(\text{trace } B) = \frac{1}{k}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} : \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \frac{2}{K} \qquad 3.10$$

$$\mathrm{varp}(\text{trace } B) = \frac{1}{k^2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} : \sum_{p,q=1}^{k}{}'' (c_p . c_q . c_p . c_q + c_p . c_q . c_q . c_p) : \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

$$= \frac{2}{k^2} \sum_{p,q=1}^{k}{}'' (a_p a_q + b_p b_q)^2 \qquad 3.11$$

Similarly

$$\mathrm{skep}(\text{trace } B) = \frac{8}{k^3} \sum_{p,q,r=1}^{k}{}'' (a_p a_q + b_p b_q)(a_p a_r + b_p b_r)(a_q a_r + b_q b_r) \quad 3.12$$

and finally

$$\mathrm{elop}(\text{trace } B) = \frac{48}{k^4} \sum_{p,q,r,s=1}^{k}{}'' (a_p a_q + b_p b_q)(a_p a_r + b_p b_r)(a_q a_s + b_q b_s)(a_r a_s + b_r b_s)$$

$$- \frac{16}{k^4} \sum_{p,q=1}^{k}{}'' (a_p a_q + b_p b_q)^4 \qquad 3.13$$

The ease of computation of the permutation cumulants of trace B higher than the first will depend on our ability to write them in a way which requires the summing of a small number of terms rather than as they are given in 3.11, 3.12 and 3.13. Since trace B is invariant with respect to the orientation of the (a, b) axes it seems desirable to compute the cumulants of B averaged over rotations of the coordinate system to which the vectors defining B are referred. In this way, the elements of these averaged cumulants of B will be invariant in the same way that trace B is invariant. It turns out that the elements of the averaged cumulants of B can all be expressed as linear combinations of seven basic invariants. These invariants are:

$$1; \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^2; \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^3; \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^4; \quad [\sum_{p=1}^{k}(a_p^2 + b_p^2)^2]^2;$$

$$\sum_{p,q=1}^{k}{}''(a_p a_q + b_p b_q)^4; \quad \text{and} \quad \sum_{p,q=1}^{k}{}''(a_p b_q - b_p a_q)^4$$

Since

$$\begin{aligned} \text{avep}\{\text{trace B}\} &= \underset{\theta}{\text{ave}}\ \text{avep}\{\text{trace B}\} \\ &= \underset{\theta}{\text{ave}}\ I_\alpha[\text{avep}\{B\}]_\alpha \\ &= I_\alpha[\underset{\theta}{\text{ave}}\ \text{avep}\{B\}]_\alpha \end{aligned}$$

and it can be shown in a similar fashion that

$$\text{varp}\{\text{trace B}\} = I_\alpha I_\beta[\underset{\theta}{\text{ave}}\ \text{varp}\{B\}]_{\alpha\beta} ,$$

$$\text{skep}\{\text{trace B}\} = I_\alpha I_\beta I_\gamma[\underset{\theta}{\text{ave}}\ \text{skep}\{B\}]_{\alpha\beta\gamma} ,$$

and

$$\text{elop}\{\text{trace B}\} = I_\alpha I_\beta I_\gamma I_\delta[\underset{\theta}{\text{ave}}\ \text{elop}\{B\}]_{\alpha\beta\gamma\delta}$$

then the cumulants of trace B can be found directly from the permutation cumulants of B averaged over all orientations of the underlying coordinate system. In this way the cumulants of trace B can be expressed as linear combinations of the seven basic invariants given above. All these invariants can be computed quickly and easily in any specific case.

3.5 Derivation of Cumulants of Trace B in Terms of a Set of Invariants

The formulas 3.11, 3.12 and 3.13 for the second, third and fourth cumulants of trace B can be expressed in the more computationally convenient way described in Section 3.4 by applying the results of Chapter 5. One can easily write these cumulants in terms of the set of invariants that are obtained for the four cumulants of B averaged over rotation.

$$\text{avep(trace B)} = \frac{2}{k}$$

$$\begin{aligned}
\text{varp(trace B)} &= \frac{2}{k^2}\sum_1^{k}{}''(a_pa_q + b_pb_q)^2 \\
&= \frac{2}{k^2}\sum_1^{k}{}''(a_p^2a_q^2 + 2\ a_pb_pa_qb_q + b_p^2b_q^2) \\
&= \frac{2}{k^2}\{(1 - \tfrac{3}{8}\sum_1^k(a_p^2 + b_p^2)^2) - \tfrac{2}{8}\sum_1^k(a_p^2 + b_p^2)^2 + (1 - \tfrac{3}{8}\sum_1^k(a_p^2 + b_p^2)^2)\} \\
&= \frac{2}{k^2}\{2 - \sum_1^k(a_p^2 + b_p^2)^2\} \qquad 3.15
\end{aligned}$$

These results can also be arrived at directly without the use of the results of Chapter 5, for

$$\begin{aligned}
\sum_1^k{}''(a_pa_q + b_pb_q)^2 &= (1 - \sum_1^k a_p^4) - 2\sum_1^k a_p^2b_p^2 + (1 - \sum_1^k b_p^4) \\
&= 2 - \sum_1^k (a_p^2 + b_p^2)^2 .
\end{aligned}$$

$$\text{skep(trace B)} = \frac{8}{k^3}\sum_1^{k}{}''(a_pa_q + b_pb_q)(a_pa_r + b_pb_r)(a_qa_r + b_qb_r)$$

$$= \frac{8}{k^3}\sum_1^{k}{}''(a_p^2a_q^2a_r^2 + 3\,a_p^2a_qb_qa_rb_r + 3\,a_pb_pa_qb_qb_r^2 + b_p^2b_q^2b_r^2)$$

$$= \frac{8}{k^3}[1 - \tfrac{9}{8}\sum_1^k(a_p^2+b_p^2)^2 + \tfrac{5}{8}\sum_1^k(a_p^2+b_p^2)^3 - \tfrac{3}{8}\sum_1^k(a_p^2+b_p^2)^2 + \tfrac{3}{8}\sum_1^k(a_p^2+b_p^2)^3$$

$$- \tfrac{3}{8}\sum_1^k(a_p^2+b_p^2)^2 + \tfrac{3}{8}\sum_1^k(a_p^2+b_p^2)^3 + 1 - \tfrac{9}{8}\sum_1^k(a_p^2+b_p^2)^2 + \tfrac{5}{8}\sum_1^k(a_p^2+b_p^2)^3]$$

$$= \frac{8}{k^3}[2 - 3\sum_1^k(a_p^2 + b_p^2)^2 + 2\sum_1^k(a_p^2 + b_p^2)^3] \qquad 3.16$$

This result also can be obtained directly for

$$\sum_1^{k}{}''(a_p^2a_q^2a_r^2 + 3\,a_p^2a_qb_qa_rb_r + 3\,a_pb_pa_qb_qb_r^2 + b_p^2b_q^2b_r^2)$$

$$= 1 - 3\sum_1^k{}''a_p^4a_q^2 - \sum_1^k a_p^6 - 3\sum_1^k{}''a_p^2a_q^2b_q^2 - 6\sum_1^k{}''a_p^3b_pa_qb_q - 3\sum_1^k a_p^4b_p^2$$

$$-3\sum_1^k{}''a_p^2b_p^2b_q^2 - 6\sum_1^k{}''a_pb_pa_qb_q^3 - 3\sum_1^k a_p^2b_p^4 + 1 - 3\sum_1^k{}''b_p^4b_q^2 - \sum_1^k b_p^6$$

$$= 2 - 3\sum_1^k a_p^4 + 2\sum_1^k a_p^6 - 3\sum_1^k a_p^2b_p^2 + 3\sum_1^k a_p^4b_p^2 + 6\sum_1^k a_p^4b_p^2 - 3\sum_1^k a_p^4b_p^2$$

$$-3\sum_1^k a_p^2b_p^2 + 3\sum_1^k a_p^2b_p^4 + 6\sum_1^k a_p^2b_p^4 - 3\sum_1^k a_p^2b_p^4 - 3\sum_1^k b_p^4 + 2\sum_1^k b_p^6$$

$$= 2 - 3\sum_1^k(a_p^4 + 2a_p^2b_p^2 + b_p^4) + 2\sum^k(a_p^6 + 3a_p^4b_p^2 + 3a_p^2b_p^4 + b_p^6)$$

$$= 2 - 3\sum_1^k(a_p^2 + b_p^2)^2 + 2\sum_1^k(a_p^2 + b_p^2)^3.$$

$$\text{alop(trace B)} = \frac{48}{k^4}\sum_1^{k}{}''(a_pa_q + b_pb_q)(a_pa_r + b_pb_r)(a_qa_s + b_qb_s)(a_ra_s + b_rb_s)$$

$$- \frac{16}{k^4}\sum_1^{k}{}''(a_pa_q + b_pb_q)^4$$

$$= \frac{48}{k^4}\sum_1{}''(a_p^2a_q^2a_r^2a_s^2 + 4\, a_p^2a_q^2a_rb_ra_sb_s + 4\, a_p^2a_qb_qa_rb_rb_s^2 + 2\, a_pb_pa_qb_qa_rb_ra_sb_s$$

$$+ 4\, a_pb_pa_qb_qb_r^2b_s^2 + b_p^2b_q^2b_r^2b_s^2) - \frac{16}{k^4}\sum^{k}{}''(a_pa_q + b_pb_q)^4$$

$$= \frac{1}{k^4}(48\sum_1^{k}{}''a_p^2a_q^2a_r^2a_s^2 - 16\sum_1^{k}{}''a_p^4a_q^4) + \frac{4}{k^4}(48\sum_1^{k}{}''a_p^2a_q^2a_rb_ra_sb_s - 16\sum_1^{k}{}''a_p^3b_pa_q^3b_q)$$

$$+ \frac{4}{k^4}(48\sum_1^{k}{}''a_p^2a_qb_qa_rb_rb_s^2 - 16\sum_1^{k}{}''a_p^2b_p^2a_q^2b_q^2) + \frac{2}{k^4}(48\sum_1^{k}{}''a_pb_pa_qb_qa_rb_ra_sb_s$$

$$- 16\sum_1^{k}{}''a_p^2b_p^2a_q^2b_q^2) + \frac{4}{k^4}(48\sum_1^{k}{}''a_pb_pa_qb_qb_r^2b_s^2 - 16\sum_1^{k}{}''a_pb_p^3a_qb_q^3)$$

$$+ \frac{1}{k^4}(48\sum_1^{k}{}''b_p^2b_q^2b_r^2b_s^2 - 16\sum_1^{k}{}''\, b_p^4b_q^4)$$

From Table 5.7 Chapter 5, one then obtains

$$\text{alop(trace B)} = \frac{1}{k^4}[(48+48) + (-108-24-24-24-108)\sum_1^{k}(a_p^2+b_p^2)^2 + (120+48+48+48+120)$$

$$\sum_1^{k}(a_p^2+b_p^2)^3 + (-\tfrac{435}{8}-\tfrac{27}{2}-\tfrac{43}{2}-\tfrac{19}{4}-\tfrac{27}{2}-\tfrac{435}{8})\sum_1^{k}(a_p^2+b_p^2)^4 + (15+6+8-2+6+15)$$

$$(\sum_1^{k}(a_p+b_p)^2)^2 + (20+14+4+8+14+20)\sum_1^{k}{}''(a_pa_q + b_pb_q)^4 + (-12-6-20+8-6-12)$$

$$\sum_1^{k}{}''(a_pb_q-b_pa_q)^4]$$

$$= \frac{16}{k^4}[6 - 18\sum_1^{k}(a_p^2+b_p^2)^2 + 24\sum_1^{k}(a_p^2+b_p^2)^3 - 12\sum_1^{k}(a_p^2+b_p^2)^4 + 3(\sum_1^{k}(a_p^2+b_p^2)^2)^2$$

$$+ 5\sum_1^{k}{}''(a_pa_q+b_pb_q)^4 - 3\sum^{k}{}''(a_pb_q-b_pa_q)^4] \qquad 3.17$$

Again this result can be obtained directly.

$$\sum_1^{k}{}'' a_p^2 a_q^2 a_r^2 a_s^2 = 1 - 6\sum_1^k a_p^4 + 8\sum_1^k a_p^6 - 6\sum_1^k a_p^8 + 3(\Sigma a_p^4)^2$$

$$\sum_1^{k}{}'' a_p^2 a_q^2 a_r b_r a_s b_s = -4\sum_1 a_p^6 b_p^2 + \sum_1 a_p^4 \sum_1 a_p^2 b_p^2 - \sum_1 a_p^2 b_p^2 + 4\sum_1 a_p^4 b_p^2 + 2\sum_1{}'' a_p^3 b_p a_q^3 b_q$$

$$\sum_1^{k}{}'' a_p^2 a_q b_q a_r b_r b_s^2 = -\sum_1^k a_p^2 b_p^2 + 2\sum_1 a_p^4 b_p^2 + (\sum_1^k a_p^2 b_p^2)^2 - 4\sum_1^k a_p^4 b_p^4 + 2\sum_1^k a_p^2 b_p^4 + 2\sum_1^k{}'' a_p^3 b_p a_q b_q^3$$

$$\sum_1^{k}{}'' a_p b_p a_q b_q a_r b_r a_s b_s = -6\sum_1^k a_p^4 b_p^4 + 3(\sum_1^k a_p^2 b_p^2)^2$$

$$\sum_1^{k}{}'' a_p b_p a_q b_q b_r^2 b_s^2 = -4\sum_1^k a_p^2 b_p^6 + \sum_1^k a_p^2 b_p^2 \sum_1^k b_p^4 - \sum_1^k a_p^2 b_p^2 + \sum_1^k a_p^2 b_p^4 + 2\sum_1^k{}'' a_p^3 b_p a_q^3 b_q$$

$$\sum_1^{k}{}'' b_p^2 b_q^2 b_r^2 b_s^2 = 1 - 6\sum_1^k b_p^4 + 8\sum_1^k b_p^6 - 6\sum_1^k b_p^8 + 3(\Sigma b_p^4)^2$$

When these formulas are substituted in the formula for elop(trace N)

and the terms are collected the result checks with equation 3.17.

Chapter 4

DERIVATION OF CUMULANTS FOR B UNDER RANDOMIZATION

4.1 Some remarks on Notation

The quantity for which Pitman found the moments corresponds to what is called $\frac{B}{A}$ in Chapter 2 in the multidimensional situation. This 'ratio' reduces to B itself when considered with respect to a suitable class of transformations with $A = I$, as described in Chapter 2. In the two dimensional case this can be written, with respect to a particular coordinate system, in the form:

$$\begin{pmatrix} \frac{1}{k}\left(\sum_{i=1}^{k} a_i\right)^2 & \frac{1}{k}\left(\sum_{i=1}^{k} a_i\right)\left(\sum_{i=1}^{k} b_i\right) \\ \frac{1}{k}\left(\sum_{i=1}^{k} a_i\right)\left(\sum_{i=1}^{k} b_i\right) & \frac{1}{k}\left(\sum_{i=1}^{k} b_i\right)^2 \end{pmatrix}$$

When the treatments are randomized over blocks the dyadic B takes the form

$$\begin{pmatrix} \frac{1}{k}\left(\sum_{i=1}^{k} \epsilon_i a_i\right)^2 & \frac{1}{k}\left(\sum_{i=1}^{k} \epsilon_i a_i\right)\left(\sum_{i=1}^{k} \epsilon_i b_i\right) \\ \frac{1}{k}\left(\sum_{i=1}^{k} \epsilon_i a_i\right)\left(\sum_{i=1}^{k} \epsilon_i b_i\right) & \frac{1}{k}\left(\sum_{i=1}^{k} \epsilon_i b_i\right)^2 \end{pmatrix}$$

Before proceeding with the cumulants averaged over randomizations, one needs some further knowledge on how to work with dyads and dyadics. It is necessary, in dealing with moments higher than the first, to find averages of 'powers' higher than the first of dyads. It is necessary to define the product of two or more dyads, and hence the square and higher powers of dyads. Since a dyad was formed by taking the product of two vectors, the square of a dyad or product of two dyads results in a four-dimensional array which can be displayed as follows:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae & af \\ ag & ah \end{pmatrix} \begin{pmatrix} be & bf \\ bg & bh \end{pmatrix}$$

$$\begin{pmatrix} ce & cf \\ cg & ch \end{pmatrix} \begin{pmatrix} de & df \\ dg & dh \end{pmatrix}$$

These four 2×2 arrays really should be thought of as a single $2 \times 2 \times 2 \times 2$ array. One might wish to think of the elements of such an array as forming a 4-dimensional cube with the elements themselves as vertices. In the same way higher powers, or more generally multiple products, can be represented. Further notations will be defined as the computation of cumulants proceeds.

4.2 Cumulants of B under Randomization

The first cumulant of B for the m-dimensional situation under randomization is

$$\underset{\substack{\text{(permutations}\\ \text{within blocks)}}}{\text{ave}^*}\left\{\begin{pmatrix} \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{i1})^2 & \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{i1})(\sum_{i=1}^{k}\epsilon_i a_{i2}) & \cdots & \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{i1})(\sum_{i=1}^{k}\epsilon_i a_{im}) \\ \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{i1})(\sum_{i=1}^{k}\epsilon_i a_{i2}) & \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{i2})^2 & \cdots & \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{i2})(\sum_{i=1}^{k}\epsilon_i a_{im}) \\ \vdots & \vdots & & \vdots \\ \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{i1})(\sum_{i=1}^{k}\epsilon_i a_{im}) & \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{i2})(\sum_{i=1}^{k}\epsilon_i a_{im}) & \cdots & \frac{1}{k}(\sum_{i=1}^{k}\epsilon_i a_{im})^2 \end{pmatrix}\right\}$$

* It is understood that in the remainder of this chapter "avep" implies average over permutations within blocks.

To make this expression more manageable, let us write it in vector notation by letting c_i represent $(a_{i1}, a_{i2}, \ldots, a_{im})$. Then

$$\mathrm{avep}(B) = \mathrm{avep}\left(\frac{1}{k}\left(\sum_{i=1}^{k}\epsilon_i c_i\right)\cdot\left(\sum_{i=1}^{k}\epsilon_i c_i\right)\right)$$

$$= \frac{1}{k}\,\mathrm{avep}\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\epsilon_i\epsilon_j c_i\cdot c_j\right)$$

$$= \frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{k} c_i\cdot c_j\ \mathrm{avep}(\epsilon_i\epsilon_j)$$

Recalling now how ϵ_i was originally defined, it is clear that

$$\mathrm{avep}(\epsilon_i\epsilon_i) = 1$$

$$\mathrm{avep}(\epsilon_i\epsilon_j) = \tfrac{1}{2}(+1) + \tfrac{1}{2}(-1) = 0 \qquad \text{for } i \neq j$$

Making use of Σ'' again to mean $\sum_{i,j=1}^{k}{}'' x_i x_j = \sum_{i=1}^{k}\sum_{j=1}^{k} x_i x_j - \sum_{i=1}^{k} x_i^2$, **we obtain**

$$\mathrm{avep}(B) = \frac{1}{k}\sum_{i=1}^{k}\mathrm{avep}\ \epsilon_i\epsilon_i c_i\cdot c_i + \frac{1}{k}\sum_{i,j=1}^{k}{}''\mathrm{avep}(\epsilon_i\epsilon_j c_i\cdot c_j)$$

$$= \frac{1}{k}\sum_{i=1}^{k} c_i\cdot c_i$$

Since, from Chapter 2, $\sum_{i=1}^{k} c_i\cdot c_i = A$ therefore $\mathrm{avep}(B) = \frac{1}{k}A$.

$$\mathrm{varp}(B) = \mathrm{avep}(B\cdot B) - \mathrm{avep}(B)\cdot\mathrm{avep}(B)$$

$$\mathrm{avep}(B\cdot B) = \mathrm{avep}\left(\frac{1}{k}\sum_{i=1}^{k}\epsilon_i c_i\cdot\sum_{i=1}^{k}\epsilon_i c_i\cdot\frac{1}{k}\sum_{i=1}^{k}\epsilon_i c_i\cdot\sum_{i=1}^{k}\epsilon_i c_i\right)$$

$$= \frac{1}{k^2}\,\mathrm{avep}\left(\sum_{p=1}^{k}\sum_{q=1}^{k}\sum_{r=1}^{k}\sum_{s=1}^{k}\epsilon_p c_p\cdot\epsilon_q c_q\cdot\epsilon_r c_r\cdot\epsilon_s c_s\right)$$

$$= \frac{1}{k^2}\,\mathrm{avep}\left(\sum_{p=1}^{k}\sum_{q=1}^{k}\sum_{r=1}^{k}\sum_{s=1}^{k}\epsilon_p\epsilon_q\epsilon_r\epsilon_s\ c_p\cdot c_q\cdot c_r\cdot c_s\right)$$

$$= \frac{1}{k^2}\left(\sum_{p=1}^{k} c_p\cdot c_p\cdot c_p\cdot c_p + \sum_{p,q=1}^{k}{}'' c_p\cdot c_p\cdot c_q\cdot c_q + \sum_{p,q=1}^{k}{}'' c_p\cdot c_q\cdot c_p\cdot c_q\right.$$

$$\left. + \sum_{p,q=1}^{k}{}'' c_p\cdot c_q\cdot c_q\cdot c_p\right)$$

$$\text{avep}(B)\cdot\text{avep}(B) = \frac{1}{k^2}\Big(\sum_{p=1}^{k} c_p\cdot c_p\Big)\cdot\Big(\sum_{p=1}^{k} c_p\cdot c_p\Big)$$

$$= \frac{1}{k^2}\Big(\sum_{p=1}^{k} c_p\cdot c_p\cdot c_p\cdot c_p + \sum_{p,q=1}^{k}{}'' c_p\cdot c_p\cdot c_q\cdot c_q\Big)$$

Thus $\text{var}_p(B) = \frac{1}{k^2}\Big(\sum_{p,q=1}^{k}{}'' c_p\cdot c_q\cdot c_p\cdot c_q + \sum_{p,q=1}^{k}{}'' c_p\cdot c_q\cdot c_q\cdot c_p\Big)$

$$\text{skep}(B) = \text{avep}((B - \text{avep}(B))\cdot(B - \text{avep}(B))\cdot(B - \text{avep}(B)))$$

$$= \text{avep}(B\cdot B\cdot B - B\cdot B\cdot\text{avep}(B) - B\cdot\text{avep}(B)\cdot B - \text{avep}(B)\cdot B\cdot B + B\cdot\text{avep}(B)\cdot\text{avep}(B)$$

$$+ \text{avep}(B)\cdot B\cdot\text{avep}(B) + \text{avep}(B)\cdot\text{avep}(B)\cdot B - \text{avep}(B)\cdot\text{avep}(B)\cdot\text{avep}(B))$$

$$\text{skep}(B) = \text{avep}(B\cdot B\cdot B) - \text{avep}(B\cdot B)\cdot\text{avep}(B) - \text{avep}(B\cdot\text{avep}(B)\cdot B) - \text{avep}(B)\cdot\text{avep}(B\cdot B)$$

$$+ 2\,\text{avep}(B)\cdot\text{avep}(B)\cdot\text{avep}(B)$$

$$\text{avep}(B\cdot B\cdot B) = \text{avep}\Big(\frac{1}{k^3}\sum_{p=1}^{k}\epsilon_p c_p\cdot\sum_{q=1}^{k}\epsilon_q c_q\cdot\sum_{r=1}^{k}\epsilon_r c_r\cdot\sum_{s=1}^{k}\epsilon_s c_s\cdot\sum_{t=1}^{k}\epsilon_t c_t\cdot\sum_{u=1}^{k}\epsilon_u c_u\Big)$$

$$= \Big(\frac{1}{k^3}\sum_{p=1}^{k} c_p\cdot c_p\cdot c_p\cdot c_p\cdot c_p\cdot c_p + \sum_{p,q=1}^{k}{}''\big(c_p\cdot c_p\cdot c_p\cdot c_p\cdot c_q\cdot c_q + c_p\cdot c_p\cdot c_p\cdot c_q\cdot c_p\cdot c_q$$

$$+ c_p\cdot c_p\cdot c_q\cdot c_p\cdot c_p\cdot c_q + c_p\cdot c_q\cdot c_p\cdot c_p\cdot c_q\cdot c_p + c_q\cdot c_p\cdot c_p\cdot c_p\cdot c_p\cdot c_q$$

$$+ c_p\cdot c_p\cdot c_p\cdot c_q\cdot c_q\cdot c_p + c_p\cdot c_p\cdot c_q\cdot c_p\cdot c_q\cdot c_p + c_p\cdot c_q\cdot c_p\cdot c_p\cdot c_q\cdot c_p$$

$$+ c_q\cdot c_p\cdot c_p\cdot c_p\cdot c_q\cdot c_p + c_p\cdot c_p\cdot c_q\cdot c_q\cdot c_p\cdot c_p + c_p\cdot c_q\cdot c_p\cdot c_q\cdot c_p\cdot c_p$$

$$+ c_q\cdot c_p\cdot c_p\cdot c_q\cdot c_p\cdot c_p + c_p\cdot c_q\cdot c_q\cdot c_p\cdot c_p\cdot c_p + c_q\cdot c_p\cdot c_q\cdot c_p\cdot c_p\cdot c_p$$

$$+ c_q\cdot c_q\cdot c_p\cdot c_p\cdot c_p\cdot c_p\big) + \sum_{p,q,r=1}^{k}{}''\big(c_p\cdot c_p\cdot c_q\cdot c_q\cdot c_r\cdot c_r + c_p\cdot c_p\cdot c_q\cdot c_r\cdot c_q\cdot c_r$$

$$+ c_p\cdot c_p\cdot c_q\cdot c_r\cdot c_r\cdot c_q + c_p\cdot c_q\cdot c_p\cdot c_q\cdot c_r\cdot c_r + c_p\cdot c_q\cdot c_q\cdot c_p\cdot c_r\cdot c_r$$

$$+ c_p\cdot c_r\cdot c_q\cdot c_q\cdot c_p\cdot c_r + c_p\cdot c_r\cdot c_q\cdot c_q\cdot c_r\cdot c_p + c_p\cdot c_q\cdot c_q\cdot c_r\cdot c_p\cdot c_r$$

$$+ c_p\cdot c_q\cdot c_q\cdot c_r\cdot c_r\cdot c_p + c_p\cdot c_q\cdot c_r\cdot c_q\cdot c_p\cdot c_r + c_p\cdot c_q\cdot c_r\cdot c_q\cdot c_r\cdot c_p$$

$$+ c_p\cdot c_r\cdot c_p\cdot c_q\cdot c_q\cdot c_r + c_p\cdot c_r\cdot c_p\cdot c_q\cdot c_r\cdot c_q + c_p\cdot c_r\cdot c_q\cdot c_p\cdot c_q\cdot c_r$$

$$+ c_p\cdot c_r\cdot c_q\cdot c_p\cdot c_r\cdot c_q\big)\Big)$$

$$\mathrm{avep}(B \bullet B) \bullet \mathrm{avep}(B) =$$

$$\frac{1}{k^3}\Big(\sum_{p=1}^{k} c_p \bullet c_p \bullet c_p \bullet c_p + \sum_{p,q=1}^{k}{}'' c_p \bullet c_p \bullet c_q \bullet c_q + \sum_{p,q=1}^{k}{}'' c_p \bullet c_q \bullet c_p \bullet c_q + \sum_{p,q=1}^{k}{}'' c_p \bullet c_q \bullet c_q \bullet c_p\Big) \bullet \sum_{p=1}^{k} c_p \bullet c_p$$

$$= \frac{1}{k^3}\Big(\sum_{p=1}^{k} c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_p + \sum_{p,q=1}^{k}{}'' c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_q \bullet c_q + \sum_{p,q=1}^{k}{}'' c_p \bullet c_p \bullet c_q \bullet c_q \bullet c_p \bullet c_p$$
$$+ \sum_{p,q=1}^{k}{}'' c_p \bullet c_p \bullet c_q \bullet c_q \bullet c_q \bullet c_q + \sum_{p,q,r=1}^{k}{}'' c_p \bullet c_p \bullet c_q \bullet c_q \bullet c_r \bullet c_r + \sum_{p,q=1}^{k}{}'' c_p \bullet c_q \bullet c_p \bullet c_q \bullet c_p \bullet c_p$$
$$+ \sum_{p,q=1}^{k}{}'' c_p \bullet c_q \bullet c_p \bullet c_q \bullet c_q \bullet c_q + \sum_{p,q,r=1}^{k}{}'' c_p \bullet c_q \bullet c_p \bullet c_q \bullet c_r \bullet c_r + \sum_{p,q=1}^{k}{}'' c_p \bullet c_q \bullet c_q \bullet c_p \bullet c_p \bullet c_p$$
$$+ \sum_{p,q=1}^{k}{}'' c_p \bullet c_q \bullet c_q \bullet c_p \bullet c_q \bullet c_q + \sum_{p,q,r=1}^{k}{}'' c_p \bullet c_q \bullet c_q \bullet c_p \bullet c_r \bullet c_r\Big)$$

Similarly

$$\mathrm{avep}(\bar{B} \bullet \mathrm{avep}(B) \bullet \bar{B}) = \frac{1}{k^3}\Big[\sum_{p=1}^{k} c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_p + \sum_{p,q=1}^{k}{}'' (c_p \bullet c_p \bullet c_q \bullet c_q \bullet c_p \bullet c_p$$
$$+ c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_q \bullet c_q + c_p \bullet c_p \bullet c_q \bullet c_q \bullet c_q \bullet c_q + c_p \bullet c_q \bullet c_p \bullet c_p \bullet c_p \bullet c_q$$
$$+ c_p \bullet c_q \bullet c_q \bullet c_q \bullet c_p \bullet c_q + c_p \bullet c_q \bullet c_p \bullet c_p \bullet c_q \bullet c_p + c_p \bullet c_q \bullet c_q \bullet c_q \bullet c_q \bullet c_p)$$
$$+ \sum_{p,q,r=1}^{k}{}'' (c_p \bullet c_p \bullet c_r \bullet c_r \bullet c_q \bullet c_q + c_p \bullet c_q \bullet c_r \bullet c_r \bullet c_p \bullet c_q + c_p \bullet c_q \bullet c_r \bullet c_r \bullet c_q \bullet c_p)\Big]$$

and

$$\mathrm{avep}(B) \bullet \mathrm{avep}(B \bullet B) = \frac{1}{k^3}\Big[\sum_{p=1}^{k} c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_p + \sum_{p,q=1}^{k}{}'' (c_q \bullet c_q \bullet c_p \bullet c_p \bullet c_p \bullet c_p$$
$$+ c_p \bullet c_p \bullet c_p \bullet c_p \bullet c_q \bullet c_q + c_q \bullet c_q \bullet c_p \bullet c_p \bullet c_q \bullet c_q + c_p \bullet c_p \bullet c_p \bullet c_q \bullet c_p \bullet c_q$$
$$+ c_q \bullet c_q \bullet c_p \bullet c_q \bullet c_p \bullet c_q + c_p \bullet c_p \bullet c_p \bullet c_q \bullet c_q \bullet c_p + c_q \bullet c_q \bullet c_p \bullet c_q \bullet c_q \bullet c_p)$$
$$+ \sum_{p,q,r=1}^{k}{}'' (c_r \bullet c_r \bullet c_p \bullet c_p \bullet c_q \bullet c_q + c_r \bullet c_r \bullet c_p \bullet c_q \bullet c_p \bullet c_q + c_r \bullet c_r \bullet c_p \bullet c_q \bullet c_q \bullet c_p)\Big]$$

$$\mathrm{avep}(B)\bullet\mathrm{avep}(B)\bullet\mathrm{avep}(B)\bullet = \frac{1}{k^3}\{(\sum_{p=1}^{k} c_p\bullet c_p)\bullet(\sum_{p=1}^{k} c_p\bullet c_p)\bullet(\sum_{p=1}^{k} c_p\bullet c_p)\}$$

$$= \frac{1}{k^3}\{\sum_{p=1}^{k} c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_p + \sum_{p,q=1}^{k}{}'' (c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_q + c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_p\bullet c_p$$

$$+ c_q\bullet c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_p) + \sum_{p,q,r=1}^{k}{}'' c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_r\bullet c_r\}$$

Hence

$$\mathrm{skep}(B) = \frac{1}{k^3}\{\sum_{p=1}^{k}(c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_p - 3\, c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_p + 2\, c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_p)$$

$$+ \sum_{p,q=1}^{k}{}'' (c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_q + c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_p\bullet c_q + c_p\bullet c_p\bullet c_q\bullet c_p\bullet c_p\bullet c_q$$

$$+ c_p\bullet c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_q + c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_q + c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_p + c_p\bullet c_p\bullet c_q\bullet c_p\bullet c_q\bullet c_p$$

$$+ c_p\bullet c_q\bullet c_p\bullet c_p\bullet c_q\bullet c_p + c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_p + c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_p\bullet c_p + c_p\bullet c_q\bullet c_p\bullet c_q\bullet c_p\bullet c_p$$

$$+ c_q\bullet c_p\bullet c_p\bullet c_q\bullet c_p\bullet c_p + c_p\bullet c_q\bullet c_q\bullet c_p\bullet c_p\bullet c_p + c_q\bullet c_p\bullet c_q\bullet c_p\bullet c_p\bullet c_p + c_q\bullet c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_p$$

$$- 3\, c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_q - 3\, c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_p\bullet c_p - 3\, c_q\bullet c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_p$$

$$- c_p\bullet c_q\bullet c_p\bullet c_q\bullet c_p\bullet c_p - c_q\bullet c_p\bullet c_q\bullet c_p\bullet c_p\bullet c_p - c_p\bullet c_q\bullet c_q\bullet c_p\bullet c_p\bullet c_p - c_q\bullet c_p\bullet c_p\bullet c_q\bullet c_p\bullet c_p$$

$$- c_p\bullet c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_q - c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_p - c_p\bullet c_q\bullet c_p\bullet c_p\bullet c_q\bullet c_p - c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_q$$

$$- c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_p\bullet c_q - c_p\bullet c_p\bullet c_q\bullet c_p\bullet c_q\bullet c_p - c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_p - c_p\bullet c_p\bullet c_q\bullet c_p\bullet c_p\bullet c_q$$

$$+ 2\, c_q\bullet c_q\bullet c_p\bullet c_p\bullet c_p\bullet c_p + 2\, c_p\bullet c_p\bullet c_p\bullet c_p\bullet c_q\bullet c_q + 2\, c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_p\bullet c_p)$$

$$+ \sum_{p,q,r=1}^{k}{}'' (c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_r\bullet c_r + c_p\bullet c_p\bullet c_q\bullet c_r\bullet c_q\bullet c_r + c_p\bullet c_p\bullet c_q\bullet c_r\bullet c_r\bullet c_q$$

$$+ c_p\bullet c_q\bullet c_p\bullet c_q\bullet c_r\bullet c_r + c_p\bullet c_q\bullet c_q\bullet c_p\bullet c_r\bullet c_r + c_p\bullet c_r\bullet c_q\bullet c_q\bullet c_p\bullet c_r + c_p\bullet c_r\bullet c_q\bullet c_q\bullet c_r\bullet c_p$$

$$+ c_p\bullet c_q\bullet c_q\bullet c_r\bullet c_p\bullet c_r + c_p\bullet c_q\bullet c_q\bullet c_r\bullet c_r\bullet c_p + c_p\bullet c_q\bullet c_r\bullet c_q\bullet c_p\bullet c_r + c_p\bullet c_q\bullet c_r\bullet c_q\bullet c_r\bullet c_p$$

$$+ c_p\bullet c_r\bullet c_p\bullet c_q\bullet c_q\bullet c_r + c_p\bullet c_r\bullet c_p\bullet c_q\bullet c_r\bullet c_q + c_p\bullet c_r\bullet c_q\bullet c_p\bullet c_q\bullet c_r + c_p\bullet c_r\bullet c_q\bullet c_p\bullet c_r\bullet c_q$$

$$- c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_r\bullet c_r - c_p\bullet c_q\bullet c_p\bullet c_q\bullet c_r\bullet c_r - c_p\bullet c_q\bullet c_q\bullet c_p\bullet c_r\bullet c_r - c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_r\bullet c_r$$

$$- c_p\bullet c_r\bullet c_q\bullet c_q\bullet c_p\bullet c_r - c_p\bullet c_r\bullet c_q\bullet c_q\bullet c_r\bullet c_p - c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_r\bullet c_r - c_p\bullet c_p\bullet c_q\bullet c_r\bullet c_q\bullet c_r$$

$$- c_p\bullet c_p\bullet c_q\bullet c_r\bullet c_r\bullet c_q + 2\, c_p\bullet c_p\bullet c_q\bullet c_q\bullet c_r\bullet c_r)$$

$$= \frac{1}{k^3} \sum_{p,q,r=1}^{k}{}'' (c_p \bullet c_q \bullet c_p \bullet c_r \bullet c_q \bullet c_r + c_p \bullet c_q \bullet c_p \bullet c_r \bullet c_r \bullet c_q + c_p \bullet c_q \bullet c_q \bullet c_r \bullet c_p \bullet c_r$$

$$+ c_p \bullet c_q \bullet c_q \bullet c_r \bullet c_r \bullet c_p + c_p \bullet c_q \bullet c_r \bullet c_p \bullet c_q \bullet c_r + c_p \bullet c_q \bullet c_r \bullet c_p \bullet c_r \bullet c_q + c_p \bullet c_q \bullet c_r \bullet c_q \bullet c_p \bullet c_r$$

$$+ c_p \bullet c_q \bullet c_r \bullet c_q \bullet c_r \bullet c_p)$$

$$= \frac{1}{k^3} \sum_{p,q,r=1}^{k}{}'' (c_p \bullet c_q + c_q \bullet c_p) \bullet (c_q \bullet c_r + c_r \bullet c_q) \bullet (c_r \bullet c_p + c_p \bullet c_r)$$

$$\mathrm{elop}(B) = \mathrm{avep}((B - \mathrm{avep}(B)) \bullet (B - \mathrm{avep}(B)) \bullet (B - \mathrm{avep}(B)) \bullet (B - \mathrm{avep}(B)))$$

$$- 3\, \mathrm{avep}((B - \mathrm{avep}(B)) \bullet (B - \mathrm{avep}(B))) \bullet \mathrm{avep}((B - \mathrm{avep}(B)) \bullet (B - \mathrm{avep}(B)))$$

$$\mathrm{ave}((B - \mathrm{avep}(B)) \bullet (B - \mathrm{avep}(B)) \bullet (B - \mathrm{avep}(B)) \bullet (B - \mathrm{avep}(B)))$$

$$= \mathrm{avep}(B \bullet B \bullet B \bullet B) - \mathrm{avep}(B \bullet B \bullet B \bullet \mathrm{avep}(B)) - \mathrm{avep}(B \bullet B \bullet \mathrm{avep}(B) \bullet B) - \mathrm{avep}(B \bullet \mathrm{avep}(B) \bullet B \bullet B)$$

$$- \mathrm{avep}(\mathrm{avep}(B) \bullet B \bullet B \bullet B) + \mathrm{avep}(B \bullet B \bullet \mathrm{avep}(B) \bullet \mathrm{avep}(B)) + \mathrm{avep}(B \bullet \mathrm{avep}(B) \bullet B \bullet \mathrm{avep}(B))$$

$$+ \mathrm{avep}(B \bullet \mathrm{avep}(B) \bullet \mathrm{avep}(B) \bullet B) + \mathrm{avep}(\mathrm{avep}(B) \bullet B \bullet B \bullet \mathrm{avep}(B))$$

$$+ \mathrm{avep}(\mathrm{avep}(B) \bullet B \bullet \mathrm{avep}(B) \bullet B) + \mathrm{avep}(\mathrm{avep}(B) \bullet \mathrm{avep}(B) \bullet B \bullet B)$$

$$- 3\, \mathrm{avep}(B) \bullet \mathrm{avep}(B) \bullet \mathrm{avep}(B) \bullet \mathrm{avep}(B)).$$

After considerable algebraic manipulation the fourth cumulant of B averaged over randomizations within blocks can be written as follows:

$$\mathrm{elop}(B) = \frac{1}{k^4}\Big(\sum_{p,q,r,s=1}^{k}{}'' [(c_p \bullet c_q + c_q \bullet c_p) \bullet (c_q \bullet c_r + c_r \bullet c_q) \bullet (c_r \bullet c_s + c_s \bullet c_r) \bullet (c_s \bullet c_p + c_p \bullet c_s)$$

$$+ (c_p \bullet c_q + c_q \bullet c_p) \bullet (c_s \bullet c_p + c_p \bullet c_s) \bullet (c_q \bullet c_r + c_r \bullet c_q) \bullet (c_r \bullet c_s + c_s \bullet c_r)$$

$$+ (c_p \bullet c_q + c_q \bullet c_p) \bullet (c_r \bullet c_s + c_s \bullet c_r) \bullet (c_s \bullet c_p + c_p \bullet c_s) \bullet (c_q \bullet c_r + c_r \bullet c_q)]$$

$$- \sum_{p,q=1}^{k}{}'' (c_p \bullet c_q + c_q \bullet c_p) \bullet (c_p \bullet c_q + c_q \bullet c_p) \bullet (c_p \bullet c_q + c_q \bullet c_p) \bullet (c_p \bullet c_q + c_q \bullet c_p)\Big) \qquad 4.1$$

For reference we repeat here the results for the lower cumulants, namely

$$\mathrm{skep}(B) = \frac{1}{k^3} \sum_{p,q,r=1}^{k}{}'' (c_p \bullet c_q + c_q \bullet c_p) \bullet (c_q \bullet c_r + c_r \bullet c_q) \bullet (c_r \bullet c_p + c_p \bullet c_r) \qquad 4.2$$

$$\mathrm{varp}(B) = \frac{1}{k^2} \sum_{p,q=1}^{k}{}'' (c_p \bullet c_q \bullet c_p \bullet c_q + c_p \bullet c_q \bullet c_q \bullet c_p) \qquad 4.3$$

$$\mathrm{avep}(B) = \frac{1}{k} \sum_{p=1}^{k} c_p \bullet c_p = \frac{1}{k} A \qquad 4.4$$

Chapter 5

DERIVATION OF THE CUMULANTS OF B AVERAGED OVER ROTATIONS

1. Some Introductory Remarks

We are interested in averaging the randomization (permutation) cumulants of B over rotations of the coordinate system in which the vectors defining B are expressed. The derivation given below is for the 2-dimensional case. Let the vector c_i be denoted by (a_i, b_i) for $1 \leq i \leq k$ where the a_i and b_i are scalars for any i. Recall now that if B is looked at in a coordinate system such that A is of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

then $\quad \sum_{i=1}^{k} a_i^2 = \sum_{i=1}^{k} b_i^2 = 1 \quad$ and $\quad \sum_{i=1}^{k} a_i b_i = 0$

Consider now an arbitrary rotation of axes in the (a, b) plane such that the new coordinates are denoted by (a', b'). Then a point (a_i, b_i) in the old coordinate system will now be given by (a_i', b_i') in the new coordinate system where

$$a_i = a_i' \cos\theta - b_i' \sin\theta$$

$$b_i = a_i' \sin\theta + b_i' \cos\theta$$

and θ is the angle between the old and the new axes.

2. Derivation of the Cumulants of B Averaged over Rotations

From Chapter 4,

$$\operatorname{avep}(B) = \frac{1}{k}\sum_{i=1}^{k} c_i * c_i = \frac{1}{k}\begin{pmatrix} \sum_{i=1}^{k} a_i^2 & \sum_{i=1}^{k} a_i b_i \\ \sum_{i=1}^{k} a_i b_i & \sum_{i=1}^{k} b_i^2 \end{pmatrix}$$

$$\underset{(\text{rotations})}{\text{ave}}\left(\sum_{i=1}^{k} a_i^2\right) = \frac{1}{2\pi}\int_0^{2\pi}\sum_{i=1}^{k}(a_i' \cos\theta - b_i' \sin\theta)^2 d\theta$$

$$= \frac{1}{2\pi}\sum_{i=1}^{k}\int_0^{2\pi}(a_i'^2 \cos^2\theta - 2a_i' b_i' \cos\theta \sin\theta + b_i'^2 \sin^2\theta) d\theta$$

$$= \frac{1}{2\pi}\left[\sum_{i=1}^{k} a_i'^2 \int_0^{2\pi}\cos^2\theta d\theta - 2\sum_{i=1}^{k} a_i' b_i' \int_0^{2\pi}\cos\theta \sin\theta \, d\theta + \sum_{i=1}^{k} b_i'^2 \int_0^{2\pi}\sin^2\theta \, d\theta\right]$$

$$= \frac{1}{2\pi}\sum_{i=1}^{k} a_i'^2 \int_0^{2\pi}(\tfrac{1}{2} + \tfrac{1}{2}\cos 2\theta) d\theta + \frac{2}{2\pi}\sum_{i=1}^{k} a_i' b_i' \int_0^{2\pi}\cos\theta d(\cos\theta) + \frac{1}{2\pi}\sum_{i=1}^{k} b_i'^2 \int_0^{2\pi}(\tfrac{1}{2} - \tfrac{1}{2}\cos 2\theta) d\theta$$

$$= \tfrac{1}{2}\sum_{i=1}^{k} a_i'^2 + \tfrac{1}{2}\sum_{i=1}^{k} b_i'^2$$

$$= \tfrac{1}{2}\sum_{i=1}^{k}(a_i'^2 + b_i'^2)$$

$$= \tfrac{1}{2}\sum_{i=1}^{k}(a_i^2 + b_i^2)$$

Also

$$\underset{(\text{rot.})}{\text{ave}}\sum_{i=1}^{k} a_i b_i = 0$$

and

$$\underset{(\text{rot.})}{\text{ave}}\sum_{i=1}^{k} b_i^2 = \tfrac{1}{2}\sum_{i=1}^{k}(a_i^2 + b_i^2)$$

$$\underset{(\mathrm{rot})}{\mathrm{ave}}\ \mathrm{avep}(B) = \begin{pmatrix} \frac{1}{2}\sum_{i=1}^{k}(a_i^2 + b_i^2) & 0 \\ 0 & \frac{1}{2}\sum_{i=1}^{k}(a_i^2 + b_i^2) \end{pmatrix}$$

$$\underset{(\mathrm{rot})}{\mathrm{ave}}\ \mathrm{varp}(B) = \underset{(\mathrm{rot})}{\mathrm{ave}}\ \frac{1}{k^2}\left[\sum_{p,q=1}^{k}{}''(c_p \cdot c_q \cdot c_p \cdot c_q + c_p \cdot c_q \cdot c_q \cdot c_p)\right]$$

$$c_p \cdot c_q \cdot c_p \cdot c_q = (a_p, b_p) \cdot (a_q, b_q) \cdot (a_p, b_p) \cdot (a_q, b_q)$$

$$= \begin{pmatrix} a_p a_q & a_p b_q \\ b_p a_q & b_p b_q \end{pmatrix} \cdot \begin{pmatrix} a_p a_q & a_p b_q \\ b_p a_q & b_p b_q \end{pmatrix}$$

This product can be written as a $2 \times 2 \times 2 \times 2$ array

$$\begin{matrix} \begin{pmatrix} a_p a_q a_p a_q & a_p a_q a_p b_q \\ a_p a_q b_p a_q & a_p a_q b_p b_q \end{pmatrix} & \begin{pmatrix} a_p b_q a_p a_q & a_p b_q a_p b_q \\ a_p b_q b_p a_q & a_p b_q b_p b_q \end{pmatrix} \\ \begin{pmatrix} b_p a_q a_p a_q & b_p a_q a_p b_q \\ b_p a_q b_p a_q & b_p a_q b_p b_q \end{pmatrix} & \begin{pmatrix} b_p b_q a_p a_q & b_p b_q a_p b_q \\ b_p b_q b_p a_q & b_p b_q b_p b_q \end{pmatrix} \end{matrix}$$

This product of two dyads contains elements of six distinct types:

(1) $a_p^2 a_q^2$; (2) $a_p^2 a_q b_q$; (3) $a_p b_q a_p b_q$; (4) $a_p^2 b_q^2$;

(5) $a_p b_p b_q^2$; (6) $b_p^2 b_q^2$

$c_p \cdot c_q \cdot c_q \cdot c_p$, the product of dyads in the variance of B, can also be written as a $2 \times 2 \times 2 \times 2$ array and the distinct elements appearing in such an array are identical with those above. We will now consider the effect on the elements (1) to (6) when the coordinate system (a, b) is rotated through an arbitrary angle θ and then the average is taken over $0 \leq \theta \leq 2\pi$.

$$\underset{(\mathrm{rot})}{\mathrm{ave}} \sum_{p,q=1}^{k}{}'' a_p^2 a_q^2 = \frac{1}{2\pi} \sum_{p,q=1}^{k}{}'' \int_0^{2\pi} (a_p' \cos\theta - b_p' \sin\theta)^2 (a_q' \cos\theta - b_q' \sin\theta)^2 d\theta$$

$$= \frac{1}{2\pi} \sum_{p,q=1}^{k}{}'' \int_0^{2\pi} (a_p'^2 \cos^2\theta - 2a_p' b_p' \cos\theta \sin\theta + b_p'^2 \sin^2\theta)$$

$$(a_q'^2 \cos^2\theta - 2a_q' b_q' \cos\theta \sin\theta + b_q'^2 \sin^2\theta) d\theta$$

$$= \frac{1}{2\pi} \sum_{p,q=1}^{k}{}'' \int_0^{2\pi} [a_p'^2 a_q'^2 \cos^4\theta - 4a_p'^2 a_q' b_q' \cos^3\theta \sin\theta$$

$$+ (2a_p'^2 b_q'^2 + 4a_p' b_p' a_q' b_q') \cos^2\theta \sin^2\theta$$

$$- 4a_p' b_p' b_q'^2 \cos\theta \sin^3\theta + b_p'^2 b_q'^2 \sin^4\theta] d\theta$$

Since

$$\frac{1}{2\pi} \int_0^{2\pi} \cos^4\theta \, d\theta = \frac{1}{2\pi} \int_0^{2\pi} \sin^4\theta \, d\theta = \frac{3}{8}$$

$$\frac{1}{2\pi} \int_0^{2\pi} \cos^2\theta \sin^2\theta \, d\theta = \frac{1}{8} \quad \text{and}$$

$$\frac{1}{2\pi} \int_0^{2\pi} \cos^3\theta \sin\theta \, d\theta = \frac{1}{2\pi} \int_0^{2\pi} \cos\theta \sin^3\theta \, d\theta = 0,$$

one obtains

$$\underset{(\mathrm{rot})}{\mathrm{ave}} \sum_{p,q=1}^{k}{}'' a_p^2 a_p^2 = \sum_{p,q=1}^{k}{}'' \left(\frac{3}{8} a_p'^2 a_q'^2 + \frac{2}{8} a_p'^2 a_q'^2 + \frac{4}{8} a_p' b_p' a_q' b_q' + \frac{3}{8} b_p'^2 b_q'^2 \right)$$

$$= \frac{3}{8} \left(\sum_{p=1}^{k} (a_p'^2 + b_p'^2) \right)^2 - \frac{4}{8} \sum_{p=1}^{k} a_p'^2 \sum_{p=1}^{k} b_p'^2 - \frac{3}{8} \sum_{p=1}^{k} (a_p'^2 + b_p'^2)^2$$

$$= \frac{3}{8} \cdot 4 - \frac{4}{8} \cdot 1 \cdot 1 - \frac{3}{8} \sum_{p=1}^{k} (a_p^2 + b_p^2)^2$$

$$= 1 - \frac{3}{8} \sum_{p=1}^{k} (a_p^2 + b_p^2)^2$$

because $a_p'^2 + b_p'^2 = a_p^2 + b_p^2$ and $\Sigma a_p^2 = \Sigma a_p'^2 = \Sigma b_p'^2 = \Sigma b_p^2 = 1$.

Similarly

$$(2) \quad \underset{(\mathrm{rot})}{\mathrm{ave}} \sum_{p,q=1}^{k}{}'' a_p^2 a_q b_q = 0$$

$$(3) \quad \underset{(\mathrm{rot})}{\mathrm{ave}} \sum_{p,q=1}^{k}{}'' a_p b_p a_q b_q = -\frac{1}{8} \sum_{p=1}^{k} (a_p^2 + b_p^2)$$

$$(4) \quad \underset{(\mathrm{rot})}{\mathrm{ave}} \sum_{p,q=1}^{k}{}'' a_p^2 b_q^2 = 1 - \frac{1}{8} \sum_{p=1}^{k} (a_p^2 + b_p^2)^2$$

$$(5) \quad \underset{(\mathrm{rot})}{\mathrm{ave}} \sum_{p,q=1}^{k}{}'' a_p b_p b_q^2 = 0$$

$$(6) \quad \underset{(\mathrm{rot})}{\mathrm{ave}} \sum_{p,q=1}^{k}{}'' b_p^2 b_q^2 = 1 - \frac{3}{8} \sum_{p=1}^{k} (a_p^2 + b_p^2)^2$$

Consider now the third cumulant as given by formula 4.2

$$\mathrm{skep}(B) = \frac{8}{k^3}\Big[\sum_{p,q,r=1}^{k}{}'' (c_p \cdot c_q + c_q \cdot c_p)(c_q \cdot c_r + c_r \cdot c_q)(c_r \cdot c_p + c_p \cdot c_r)$$

$$= \frac{8}{k^3}\Big[\sum_{p,q,r=1}^{k}{}'' (c_p \cdot c_q \cdot c_p \cdot c_r \cdot c_q \cdot c_r + c_p \cdot c_q \cdot c_p \cdot c_r \cdot c_r \cdot c_q + c_p \cdot c_q \cdot c_r \cdot c_p \cdot c_q \cdot c_r$$

$$+ c_p \cdot c_q \cdot c_r \cdot c_p \cdot c_r \cdot c_q + c_p \cdot c_q \cdot c_q \cdot c_r \cdot c_p \cdot c_r + c_p \cdot c_q \cdot c_q \cdot c_r \cdot c_r \cdot c_p$$

$$+ c_p \cdot c_q \cdot c_r \cdot c_q \cdot c_p \cdot c_r + c_p \cdot c_q \cdot c_r \cdot c_q \cdot c_r \cdot c_p)\Big]$$

By a similar procedure as was used for the average over rotations of the variance of B it can be shown that all the above terms in the skewness of B contain only elements of the following types:

(1) $\sum\limits_{p,q,r=1}^{k}{}'' a_p^2 a_q^2 a_r^2$ (6) $\sum\limits_{p,q,r=1}^{k}{}'' a_p^2 b_q^2 b_r^2$

(2) $\sum\limits_{p,q,r=1}^{k}{}'' a_p^2 a_q^2 a_r b_r$ (7) $\sum\limits_{p,q,r=1}^{k}{}'' a_p b_p a_q b_q a_r b_r$

(3) $\sum\limits_{p,q,r=1}^{k}{}'' a_p^2 a_q^2 b_r^2$ (8) $\sum\limits_{p,q,r=1}^{k}{}'' a_p b_p a_q b_q b_r^2$

(4) $\sum\limits_{p,q,r=1}^{k}{}'' a_p^2 a_q b_q a_r b_r$ (9) $\sum\limits_{p,q,r=1}^{k}{}'' a_p b_p b_q^2 b_r^2$

(5) $\sum\limits_{p,q,r=1}^{k}{}'' a_p^2 a_q b_q b_r^2$ (10) $\sum\limits_{p,q,r=1}^{k}{}'' b_p^2 b_q^2 b_r^2$

Instead of doing the rotation on these elements directly, as was done for the first two cumulants, an alternative procedure will be used which eliminates some of the drudgery of the algebra. This procedure proves to be very satisfactory for the averaging of the fourth cumulant.

The alternative procedure makes use of the fact that after averaging over rotations the elements are given by linear combinations of the invariants. By finding all the invariants in advance, one can easily find the coefficients which express each of the elements as such linear combinations.

Since

$$a = a' \cos\theta - b' \sin\theta \text{ and}$$

$$b = a' \sin\theta + b' \cos\theta$$

then

$$a' = a \cos\theta + b \sin\theta \text{ and}$$

$$b' = -a \sin\theta + b \cos\theta,$$

so an infinitesimal change $\Delta\theta$ in θ will change a' by an amount $b'\Delta\theta$. Hence an infinitesimal change $\Delta\theta$ in θ will change

$\overset{k}{\underset{p,q,r=1}{\Sigma''}} a_p^2 a_q^2 a_r^2$ by an amount $6\Delta\theta \overset{k}{\underset{p,q,r=1}{\Sigma''}} a_p b_p a_q^2 a_r^2$. A two-way table can be constructed to show the effect of an infinitesimal change in θ on each of the ten elements pertaining to the skewness.

Table 5.1

Coefficients of the Linear Combinations of the Elements of skep b after an Infinitesimal Change in θ

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) $\Sigma'' a_p^2 a_q^2 a_r^2$	-	6								
(2) $a_p^2 a_q^2 a_r b_r$	-1	-	1	4						
(3) $a_p^2 a_q^2 b_r^2$		-2	-	4						
(4) $a_p^2 a_q b_q a_r b_r$		-2		-	2		2			
(5) $a_p^2 a_q b_q b_r^2$			-1	-2	-	1		2		
(6) $a_p^2 b_q^2 b_r^2$					-4	-			2	
(7) $a_p b_p a_q b_q a_r b_r$				-3			-	3		
(8) $a_p b_p a_q b_q b_r^2$					-2		-2	-	-2	
(9) $a_p b_p b_q^2 b_r^2$						-1		-4	-	1
(10) $b_p^2 b_q^2 b_r^2$									-6	-

The numbers in Table 5.1 refer to the coefficients in the linear combinations obtained by making an infinitesimal change in θ. For example, a change $\Delta\theta$ in θ changes $\overset{k}{\underset{p,q,r=1}{\Sigma''}} a_p b_p b_q^2 b_r^2$ by

$(-\overset{k}{\underset{p,q,r=1}{\Sigma''}} a_p^2 b_q^2 b_r^2 - 4 \overset{k}{\underset{p,q,r=1}{\Sigma''}} a_p b_p a_q b_q b_r^2 + \overset{k}{\underset{p,q,r=1}{\Sigma''}} b_p^2 b_q^2 b_r^2)\Delta\theta$. To find all the invariants under rotation one only needs to find those combinations of (1) through (10) which are unaltered by an infinitesimal change in θ. These are:

$$(a)\quad = \sum_{p,q,r=1}^{k}{}'' (a_p^2 a_q^2 a_r^2 + 2\, a_p^2 a_q b_q a_r b_r + a_p^2 a_q^2 b_r^2 + a_p^2 b_q^2 b_r^2 + 2\, a_p b_p a_q b_q b_r^2 + b_p^2 b_q^2 b_r^2)$$

$$= \sum_{p,q,r=1}^{k}{}'' [a_p^2(a_q^2 a_r^2 + 2\, a_q b_q a_r b_r + b_q^2 b_r^2) + b_p^2(a_q^2 a_r^2 + 2\, a_q b_q a_r b_r + b_q^2 b_r^2)$$

$$= \sum_{p,q,r=1}^{k}{}'' (a_p^2 + b_p^2)(a_q a_r + b_q b_r)^2$$

$$= \sum_{p=1}^{k} (a_p^2 + b_p^2) \sum_{p,q=1}^{k}{}'' (a_p a_q + b_p b_q)^2 - 2\sum_{p,q=1}^{k}{}'' (a_p^2 + b_p^2)(a_p a_q + b_p b_q)^2$$

But

$$\sum_{p,q=1}^{k}{}'' (a_p a_q + b_p b_q)^2 = [(\sum_{p=1}^{k} a_p^2)^2 - \sum_{p=1}^{k} a_p^4] + 2[(\sum_{p=1}^{k} a_p b_p)^2 - \Sigma a_p^2 b_p^2] + [(\Sigma b_p^2)^2 - \Sigma b_p^4]$$

$$= 2 - \sum_{p=1}^{k} (a_p^2 + b_p^2)^2$$

and

$$\sum_{p,q=1}^{k}{}'' (a_p^2 + b_p^2)(a_p a_q + b_p b_q)^2$$

$$= \sum_{p,q=1}^{k}{}'' (a_p^4 a_q^2 + 2a_p^3 b_p a_q b_q + a_p^2 b_p^2 b_q^2 + a_p^2 b_p^2 a_q^2 + 2a_p b_p^3 a_q b_q + b_p^4 a_q^2)$$

$$= \sum_{p=1}^{k} (a_p^4 - a_p^6 + 2a_p^4 b_p^2 + a_p^2 b_p^2 - a_p^2 b_p^4 + a_p^2 b_p^2 - a_p^4 b_p^2 - 2a_p^2 b_p^4 + b_p^4 - b_p^6)$$

$$= \sum_{p=1}^{k} (a_p^2 + b_p^2)^2 - \Sigma(a_p^2 + b_p^2)^3.$$

Thus

$$(a) = 4 - 2\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 - 2\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 + 2\sum_{p=1}^{k}(a_p^2 + b_p^2)^3$$

$$= 4 - 4\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 + 2\sum_{p=1}^{k}(a_p^2 + b_p^2)^3.$$

Similarly

$$(b) = \sum_{p,q,r=1}^{k}{}''(a_p^2a_q^2a_r^2 + 3a_p^2a_q^2b_r^2 + 3a_p^2b_q^2b_r^2 + b_p^2b_q^2b_r^2) = \sum_{p,q,r=1}^{k}{}''(a_p^2+b_p^2)(a_q^2+b_q^2)(a_r^2+b_r^2)$$

$$= 8 - 6\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 + 2\sum_{p=1}^{k}(a_p^2 + b_p^2)^3$$

and

$$(c) = \sum_{p,q,r=1}^{k}{}''(a_p^2a_qb_qa_rb_r - a_p^2a_q^2b_r^2 + a_p^2a_qb_qb_r^2 - a_p^2b_q^2b_r^2)$$

$$= \sum_{p,q,r=1}^{k}{}''(a_pb_pa_qb_q - a_p^2b_q^2)(a_r^2 + b_r^2) = -2 + \sum_{p=1}^{k}(a_p^2 + b_p^2)^2.$$

These three invariants (a), (b) and (c) are such that no other linear combination of (1) through (10) can be found independent of them. Furthermore, these three invariants are independent of each other and are all expressible in the form

$$\alpha + \beta\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 + \gamma\sum_{p=1}^{k}(a_p^2 + b_p^2)^3$$

where α, β and γ are constants. Consequently

$$1, \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^2 \quad \text{and} \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^3$$

are three convenient invariants and all elements of the third cumulants of B (when averaged over all rotations of the coordinate system) must be expressible as a linear combination of 1, $\sum_{p=1}^{k}(a_p^2 + b_p^2)^2$ and $\sum_{p=1}^{k}(a_p^2 + b_p^2)^3$.

By choosing some simple sets of vectors one can easily compute the values of (1) through (10). This method will now be outlined. Let us choose the following three sets of vectors:

$$\begin{array}{ccc} (\alpha) & (\beta) & (\gamma) \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 0 \\ 0 & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \end{pmatrix} \end{array}$$

for the set α

$$\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 = 2, \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^3 = 2$$

for the set β

$$\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 = 1, \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^3 = \frac{1}{2}$$

for the set γ

$$\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 = \frac{2}{3}, \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^3 = \frac{2}{9}$$

Any third cumulant element averaged over rotations is of the form

$$\underset{(\text{rot.})}{\text{ave}}(L) = x + y\,\Sigma(a_p^2 + b_p^2)^2 + z\,\Sigma(a_p^2 + b_p^2)^3$$

where L is an arbitrary expression of sixth degree in a and b. Let α_o, β_o, γ_o be the values of an arbitrary quantity L averaged over rotations for the sets α, β and γ respectively. Then

$$\begin{aligned} \alpha_o &= x + 2y + 2z \\ \beta_o &= x + y + \tfrac{1}{2}z \\ \gamma_o &= x + \tfrac{2}{3}y + \tfrac{2}{9}z \end{aligned}$$

and, solving these equations,

$$\underset{(\text{rot.})}{\text{ave}(L)} = \left(\tfrac{1}{2}\alpha_o - 4\beta_o + \tfrac{9}{2}\gamma_o\right) + \left(-\tfrac{5}{4}\alpha_o + 8\beta_o - \tfrac{27}{4}\gamma_o\right)\sum_{p=1}^{k}(a_p^2 + b_p^2)^2$$

$$+ \left(\tfrac{3}{2}\alpha_o - 6\beta_o + \tfrac{9}{2}\gamma_o\right)\sum_{p=1}^{k}(a_p^2 + b_p^2)^3$$

A table of α_o, β_o and γ_o for (1) through (10) is shown below as Table 5.2. The values of the coefficients x, y and z in the expression for $\underset{(\text{rot.})}{\text{ave}(L)}$ may then be written down, using this table. These values are given in Table 5.3.

Table 5.2

Values of the elements in skep B averaged over θ for the vectors α, β and γ

	α_o	β_o	γ_o
(1)	0	3/16	7/18
(2)	0	0	0
(3)	0	7/16	11/18
(4)	0	-1/16	-1/18
(5)	0	0	0
(6)	0	7/16	11/18
(7)	0	0	0
(8)	0	-1/16	-1/18
(9)	0	0	0
(10)	0	3/16	7/18

Table 5.3

Coefficients in the linear combinations expressing the elements of ave skep B in terms of three convenient invariants.
(rot)

	1	$\Sigma(a_p^2 + b_p^2)^2$	$\Sigma(a_p^2 + b_p^2)^3$
(1)	1	-9/8	5/8
(2)	0	0	0
(3)	1	-5/8	1/8
(4)	0	-1/8	1/8
(5)	0	0	0
(6)	1	-5/8	1/8
(7)	0	0	0
(8)	0	-1/8	1/8
(9)	0	0	0
(10)	1	-9/8	5/8

The fourth cumulant of B contains product of dyads of the form $c_p \cdot c_q \cdot c_p \cdot c_r \cdot c_q \cdot c_s \cdot c_r \cdot c_s$ and of the form $c_p \cdot c_q \cdot c_p \cdot c_q \cdot c_p \cdot c_q \cdot c_p \cdot c_q$. To find the average over rotations of elop B we first need to find a maximal set of independent invariants. It will then be possible to find that linear combination of these invariants which results, after rotation, from each of the distinct elements in $c_p \cdot c_q \cdot \ldots \cdot c_s$ and $c_p \cdot c_q \cdot \ldots \cdot c_q$. The distinct elements are

1)	$\sum_{p,q,r,s=1}^{k} a_p^2 a_q^2 a_r^2 a_s^2$	16)	$\sum_{p,q=1}^{k} a_p^4 a_q^4$
2)	$a_p^2 a_q^2 a_r^2 a_s b_s$	17)	$a_p^4 a_q^3 b_q$
3)	$a_p^2 a_q^2 a_r^2 b_s^2$	18)	$a_p^4 a_q^2 b_q^2$
4)	$a_p^2 a_q^2 a_r b_r a_s b_s$	19)	$a_p^4 a_q b_q^3$
5)	$a_p^2 a_q^2 a_r b_r b_s^2$	20)	$a_p^4 b_q^4$
6)	$a_p^2 a_q^2 b_r^2 b_s^2$	31)	$a_p^3 b_p a_q^3 b_q$
7)	$a_p^2 a_q b_q a_r b_r a_s b_s$	22)	$a_p^3 b_p a_q^2 b_q^2$
8)	$a_p^2 a_q b_q a_r b_r b_s^2$	23)	$a_p^3 b_p a_q b_q^3$
9)	$a_p^2 a_q b_q b_r^2 b_s^2$	24)	$a_p^3 b_p b_q^4$
10)	$a_p^2 b_q^2 b_r^2 b_s^2$	25)	$a_p^2 b_p^2 a_q^2 b_q^2$
11)	$a_p b_p a_q b_q a_r b_r a_s b_s$	26)	$a_p^2 b_p^2 a_q b_q^3$
12)	$a_p b_p a_q b_q a_r b_r b_s^2$	27)	$a_p^2 b_p^2 b_q^4$
13)	$a_p b_p a_q b_q b_r^2 b_s^2$	28)	$a_p b_p^3 a_q b_q^3$
14)	$a_p b_p b_q^2 b_r^2 b_s^2$	29)	$a_p b_p^3 b_q^4$
15)	$b_p^2 b_q^2 b_r^2 b_s^2$	30)	$b_p^4 b_q^4$

The coefficients relating 1) through 15) and 16) through 30) with the result of an infinitesimal change in θ are given in Table 5.4

Table 5.4

Coefficients of the linear combinations of the elements of elop **B** after an infinitesimal change in θ

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-	8													
2	-1	-	1	6											
3		-2	-		6										
4		-2		-	-2		4								
5			-1	-2	-	1		4							
6					-4	-			4						
7				-3			-	3			2				
8					-2		-2	-	2			2			
9						-1		-4	-	1			2		
10									-6	-				2	
11							-4				-	4			
12								-3			-2	-	3		
13									-2			-4	-	2	
14										-1			-6	-	1
15														-8	-

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
16	-	8													
17	-1	-	3			4									
18		-2	-	2			4								
19			-3	-	1			4							
20				-4	-				4						
21		-2				-	6								
22			-1			-2	-	2		[illegible]					
23				-1			-3	-	1		3				
24					-1			-4	-			3			
25							-4			-	4				
26								-2		-3	-	1	2		
27									[illegible]		-4	-		2	
28												-6	-	2	
29												-3	-4	-	1
30														-8	-

From the part of Table 4.4 involving (1) through (15) the following set of independent invariants can be deduced:

(A) = (1) + 4(4) + 2(6) + 4(11) + 4(13) + (15)

(B) = (1) + 4(3) + 6(6) + 4(10) + (15)

(C) = (6) - 2(8) + (11)

By rearrangement and some algebra (A), (B) and (C) can be written as

$$A = 4 - 8\sum_{p=1}^{k}(a_p^2 + b_p^2)^2 + 8\sum_{p=1}^{k}(a_p^2 + b_p^2)^3 - 4\sum_{p=1}^{k}(a_p^2 + b_p^2)^4$$

$$+ [\sum_{p=1}^{k}(a_p^2 + b_p^2)^2]^2 + 2\sum_{p,q=1}^{k}{}''(a_p a_q + b_p b_q)^4$$

$$B = 16 - 6\sum_{p=1}^{k}(a_p^2 + b_p^2)^4 + 16\sum_{p=1}^{k}(a_p^2 + b_p^2)^3 - 24\sum_{p=1}^{k}(a_p^2 + b_p^2)^2$$

$$+ 3[\sum_{p=1}^{k}(a_p^2 + b_p^2)^2]^2$$

$$C = 1 - \sum_{p=1}^{k}(a_p^2 + b_p^2)^2 + \frac{1}{2}\sum_{p,q=1}^{k}{}''(a_p b_q - b_p a_q)^4$$

From the table involving (16) through (30) the following set of independent invariants can be deduced

(A') = (16) + 4(21) + 6(25) + 4(28) + 30

(B') = (16) + 4(18) + 2(20) + 4(25) + 4(27) + (30)

(C') = (20) - 4(23) + 3(25)

These can also be written as

$$(A') = \sum_{p,q=1}^{k}{}''(a_p a_q + b_p b_q)^4$$

$$(B') = [\sum_{p=1}^{k}(a_p^2 + b_p^2)^2]^2 - \Sigma(a_p^2 + b_p^2)^4$$

$$(C') = \frac{1}{2}\sum_{p,q=1}^{k}{}''(a_p b_q - b_p a_q)^4$$

From a careful study of these invariants it soon becomes apparent that a convenient maximal set of independent invariants, as linear combinations of which we may express the average over rotations of the fourth cumulant of B, are 1,

$$\sum_{p=1}^{k}(a_p^2 + b_p^2)^2, \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^3, \quad \sum_{p=1}^{k}(a_p^2 + b_p^2)^4,$$

$$[\sum_{p=1}^{k}(a_p^2 + b_p^2)^2]^2, \quad \sum_{p,q=1}^{k}{}''(a_p a_q + b_p b_q)^4, \quad \sum_{p,q=1}^{k}{}''(a_p b_q - b_p a_q)^4 .$$

The average values of the terms (16) through (30) for all rotations of the coordinate system only involve the linear combinations of three invariants (A'), (B'), (C'). A table of the coefficients of (A'), (B'), (C') for each of the fifteen is now given as Table 5.5.

Table 5.5

Coefficients of the linear combinations of (A'), (B'), (C') for the elements (1) through (15) averaged over rotations. (Each entry must be divided by 128)

	(A')	(B')	(C')		(A')	(B')	(C')
(16)	20	15	-12	(23)	0	3	-8
(17)	0	0	0	(24)	0	0	0
(18)	-4	9	-4	(25)	4	-1	4
(19)	0	0	0	(26)	0	0	0
(20)	-12	15	20	(27)	-4	9	-4
(21)	8	-3	0	(28)	8	-3	0
(22)	0	0	0	(29)	0	0	0
				(30)	20	15	-12

The average values of the terms (1) through (15) for all rotations of the coordinate system involve all of the seven invariants and are given in Table 5.6. Both Tables 5.5 and 5.6 were calculated by direct rotation methods as given earlier in this section for the second cumulant and also by the method outlined for the third cumulant.

Table 5.6

Coefficients of the linear combinations expressing the elements (1) through (15) in terms of a set of invariants. (Each entry must be divided by 128.)

	1	$\Sigma(a_p^2+b_p^2)^2$	$\Sigma(a_p^2+b_p^2)^3$	$\Sigma(a_p^2+b_p^2)^4$	$[\Sigma(a_p^2+b_p^2)^2]^2$	$\Sigma(a_pa_q+b_pb_q)^4$	$\Sigma(a_pb_q-b_pa_q)^4$
(1)	128	-288	320	-150	45	60	-36
(2)	0	0	0	0	0	0	0
(3)	128	-192	128	-42	27	-12	-12
(4)	0	-16	32	-18	3	12	-4
(5)	0	0	0	0	0	0	0
(6)	128	-160	64	-22	13	-4	28
(7)	0	0	0	0	0	0	0
(8)	0	-16	32	-14	5	4	-12
(9)	0	0	0	0	0	0	0
(10)	128	-192	128	-42	27	-12	-12
(11)	0	0	0	-6	-3	12	12
(12)	0	0	0	0	0	0	0
(13)	0	-16	32	-18	3	12	-4
(14)	0	0	0	0	0	0	0
(15)	128	-288	320	-150	45	60	-36

By combining the elements (1) through (15) with elements (16) through (30) according to formula 4.4, we arrive at Table 5.7 which has 21 entries. Some one of these 21 terms is equivalent to each of the 256 terms which arise in the 2^8 array defining the fourth cumulant of B averaged over rotations of the coordinate system.

Table 5.7

Coefficients of the linear combinations expressing the elements of ave elop B in terms of seven convenient invariants (rot.)

		1	$\sum_{p=1}^{k}(a_p^2+b_p^2)^2$	$\sum_{p=1}^{k}(a_p^2+b_p^2)^3$	$\sum_{p=1}^{k}(a_p^2+b_p^2)^4$	$\left[\sum_{p=1}^{k}(a_p^2+b_p^2)^2\right]^2$	$\sum_{p,q=1}^{k}(a_p a_q+b_p b_q)^4$	$\sum_{p,q=1}^{k}(a_p b_q-b_p a_q)^4$
$-16\sum_{p,q=1}^{k}{}'' a_p^4 a_q^4$	$+48\sum_{p,q,r,s=1}^{k}{}'' a_p^2 a_q^2 a_r^2 a_s^2$	48	-108	120	$-\frac{435}{8}$	15	20	-12
$a_p^4 a_q^3 b_q$	$a_p^2 a_q^2 a_r^2 a_s b_s$	0	0	0	0	0	0	0
$a_p^4 a_q^2 b_q^2$	$a_p^2 a_q^2 a_r^2 b_s^2$	48	-72	48	$-\frac{117}{8}$	9	-4	-4
$a_p^4 a_q^2 b_q^2$	$a_p^2 a_q^2 a_r b_r a_s b_s$	0	-6	12	$-\frac{45}{8}$	0	5	-1
$a_p^4 a_q b_q^3$	$a_p^2 a_q^2 a_r b_r b_s^2$	0	0	0	0	0	0	0
$a_p^4 b_q^4$	$a_p^2 a_q^2 b_r^2 b_s^2$	48	-60	24	$-\frac{51}{8}$	3	0	8
$a_p^3 b_p a_q^3 b_q$	$a_p^2 a_q b_q a_r^2 a_s b_s$	0	-6	12	$-\frac{57}{8}$	$\frac{3}{2}$	$\frac{7}{2}$	$-\frac{3}{2}$
$a_p^3 b_p a_q^2 b_q^2$	$a_p^2 a_q b_q a_r b_r a_s b_s$	0	0	0	0	0	0	0
$a_p^3 b_p a_q^2 b_q^2$	$a_p^2 a_q b_q a_r^2 b_s^2$	0	0	0	0	0	0	0
$a_p^3 b_p a_q b_q^3$	$a_p^2 a_q b_q a_r b_r b_s^2$	0	-6	12	$-\frac{39}{8}$	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{7}{2}$
$a_p^3 b_p b_q^4$	$a_p^2 a_q b_q b_r^2 b_s^2$	0	0	0	0	0	0	0
$a_p^2 b_p^2 a_q^2 b_q^2$	$a_p^2 b_q^2 a_r b_r a_s b_s$	0	-6	12	$-\frac{43}{8}$	2	1	-5
$a_p^2 b_p^2 a_q^2 b_q^2$	$a_p^2 b_q^2 a_r^2 b_s^2$	48	-60	24	$-\frac{67}{8}$	5	-2	10
$a_p^2 b_p^2 a_q b_q^3$	$a_p^2 b_q^2 a_r b_r b_s^2$	0	0	0	0	0	0	0
$a_p^2 b_p^2 b_q^4$	$a_p^2 b_q^2 b_r^2 b_s^2$	48	-72	48	$-\frac{117}{8}$	9	-4	-4
$a_p^2 b_p^2 a_q^2 b_q^2$	$a_p b_p a_q b_q a_r b_r a_s b_s$	0	0	0	$-\frac{19}{8}$	-1	4	4
$a_p^2 b_p^2 a_q b_q^3$	$a_p b_p a_q b_q a_r b_r b_s^2$	0	0	0	0	0	0	0
$a_p^2 b_p^2 b_q^4$	$a_p b_p a_q b_q b_r^2 b_s^2$	0	-6	12	$-\frac{45}{8}$	0	5	-1
$a_p b_p^3 a_q b_q^3$	$a_p b_p b_q^2 a_r b_r b_s^2$	0	-6	12	$-\frac{57}{8}$	$\frac{3}{2}$	$\frac{7}{2}$	$-\frac{3}{2}$
$a_p b_p^3 b_q^4$	$a_p b_p b_q^2 b_r^2 b_s^2$	0	0	0	0	0	0	0
$b_p^4 b_q^4$	$b_p^2 b_q^2 b_r^2 b_s^2$	48	-108	120	$-\frac{435}{8}$	15	20	-12

Chapter 6

NORMAL THEORY FOR TWO TREATMENTS

6.1 Introduction: t^2 in One Dimension

A standard method for testing whether the mean of a sample has a certain specified value when the population variance is unknown is to make use of the t distribution. Suppose that $x_1, x_2, \ldots, x_k$ is a sample of k independent observations from a normal distribution with unknown mean μ and standard deviation σ. To test the hypothesis that the mean of the population has some specified value, say μ_0, one forms the ratio

$$\frac{\bar{x} - \mu_0}{s/\sqrt{k}} \qquad 6.1$$

where $\bar{x} = \sum_{i=1}^{k} x_i$ and $s = \sqrt{\frac{1}{k-1}\sum_{i=1}^{k}(x_i - \bar{x})^2}$

This ratio has Student's distribution with $(k-1)$ degrees of freedom.

In the same way, a test for the difference between two population means makes use of the assumption that the populations are independently normally distributed with unknown means μ_1 and μ_2. If the variances are also unknown but can be assumed equal then a statistic for testing whether the means are equal makes use of a pooled estimate of the population variance defined as

$$s^2 = \frac{s_1^2(k_1 - 1) + s_2^2(k_2 - 1)}{k_1 + k_2 - 2}$$

where

(i) $$s_i^2 = \frac{1}{k_i - 1}\sum_{j=1}^{k_i}(x_{ij} - \bar{x}_i)^2 \qquad i = 1, 2$$

(ii) $x_{11}, x_{12}, \ldots, x_{1k_1}$ and $x_{21}, x_{22}, \ldots, x_{2k_2}$ are the samples from the populations

(iii) $$\bar{x}_i = \frac{1}{k_i} \sum_{j=1}^{k_i} x_{ij} \qquad i = 1, 2$$

The statistic used is

$$\frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{k_1} + \frac{1}{k_2}}} \qquad 6.2$$

The ratio is distributed as t with $(k_1 + k_2 - 2)$ degrees of freedom and the test is made in the same way as the single sample test given above.

If one is not able to make the assumption of equal population variances then the problem is more complex. The separate estimates s_1^2 and s_2^2 for the population variances are used in the ratio

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{k_1} + \frac{s_2^2}{k_2}}} \qquad 6.3$$

This ratio is not distributed as t with $(k_1 + k_2 - 2)$ degrees of freedom as was the case with equal variances because in general the denominator is not proportional to a χ^2 with $(k_1 + k_2 - 2)$ degrees of freedom. However, this distribution has been studied by Welch [16] and significance points have been tabled. Furthermore, Welch [17] showed that the percentage points can be approximated by the percentage points of the t distribution with f degrees of freedom where

$$f = \frac{(k_2 - 1)\, s_1^4/k_1^2 + (k_1 - 1)\, s_2^4/k_2^2}{(k_1 - 1)(k_2 - 1)(s_1^2/k_1 + s_2^2/k_2)^2}$$

These tests make use of the assumption that the underlying populations are normally distributed but as mentioned in Chapter 1, they are remarkably insensitive to deviations from normality.

6.2 Introduction: Hotelling's T^2

Proceeding now to the multivariate problem, we assume

$\{(x_{11}, x_{21}),\ (x_{12}, x_{22}),\ \ldots,\ (x_{1k}, x_{2k})\}$

are pairs of observations drawn independently from a bivariate normal distribution with parameters μ_1, μ_2, σ_1^2, σ_2^2 and ρ. The sample means $(\bar{x}_1, \bar{x}_2)$ are distributed according to a bivariate normal distribution with parameters μ_1, μ_2, σ_1^2/k, σ_2^2/k and ρ. These sample means are independent of s_1^2, s_2^2 and r, the estimates of σ_1^2, σ_2^2, and ρ respectively which are defined as

$$s_i^2 = \frac{1}{k-1} \sum_{j=1}^{k} (x_{ij} - \bar{x}_i)^2 \qquad i = 1, 2.$$

$$r = \frac{\sum_{j=1}^{k} (x_{1j} - \bar{x}_1)(x_{2j} - \bar{x}_2)}{\sqrt{\sum_{j=1}^{k} (x_{1j} - \bar{x}_1)^2 \sum_{j=1}^{k} (x_{2j} - \bar{x}_2)^2}} \qquad 6.3$$

The quantity

$$\frac{1}{1-\rho^2}\left[\frac{(\bar{x}_1 - \mu_1)^2}{\sigma_1^2/k} - 2\rho\, \frac{(\bar{x}_1 - \mu_1)}{\sigma_1/\sqrt{k}}\, \frac{(\bar{x}_2 - \mu_2)}{\sigma_2/\sqrt{k}} + \frac{(\bar{x}_2 - \mu_2)^2}{\sigma_2^2/k}\right] \qquad 6.4$$

is distributed as χ^2 with 2 degrees of freedom. Replacing σ_1^2, σ_2^2 and ρ by the estimates s_1^2, s_2^2 and r respectively, expression 6.4 is replaced by

$$T^2 = \frac{1}{1-r^2}\left[\frac{(\bar{x}_1-\mu_1)^2}{s_1^2/k} - 2r\,\frac{(\bar{x}_1-\mu_1)}{s_1/\sqrt{k}}\,\frac{(\bar{x}_2-\mu_2)}{s_2/\sqrt{k}} + \frac{(\bar{x}_2-\mu_2)^2}{s_2^2/k}\right] \qquad 6.5$$

When (x_{1j}, x_{2j}), $1 \leq j \leq k$, are independent samples from a bivariate normal, then T^2 is distributed as

$$2\,\frac{(k-1)}{(k-2)}\,F_{2,k-2} \qquad 6.6$$

To test the hypothesis at the $\alpha\%$ level that μ_1 and μ_2 have specified values μ_{10} and μ_{20} say, against the alternative that $\mu_1 \neq \mu_{10}$ and $\mu_2 \neq \mu_{20}$ one computes T^2 with μ_1 and μ_2 replaced by μ_{10} and μ_{20} and compares the result with the value of 6.6 using the table of F with 2 and $(k-2)$ degrees of freedom at the $\alpha\%$ point.

The two sample problem in more than one dimension follows a similar procedure. This will be illustrated by testing that the difference between the means of two bivariate normal distributions might reasonably be zero. Let (x_{1i}, x_{2i}) and (y_{1j}, y_{2j}), $1 \leq i \leq k_1$ and $1 \leq j \leq k_2$, be independent samples from bivariate normal populations with parameters μ_{x_1}, μ_{x_2}, $\sigma^2_{x_1}$, $\sigma^2_{x_2}$, $\rho_{x_1x_2}$ and μ_{y_1}, μ_{y_1}, $\sigma^2_{y_1}$, $\sigma^2_{y_2}$, $\rho_{y_1y_2}$. Let the estimates of these parameters be denoted by $\bar{x}_1$, $\bar{x}_2$, $s^2_{x_1}$, $s^2_{x_2}$, $r_{x_1x_2}$ and $\bar{y}_1$, $\bar{y}_2$, $s^2_{y_1}$, $s^2_{y_2}$, $r_{y_1y_2}$. Assuming that $\sigma^2_{x_1} = \sigma^2_{y_1} = \sigma^2_1$, $\sigma^2_{x_2} = \sigma^2_{y_2} = \sigma^2_2$ and $\rho_{x_1x_2} = \rho_{y_1y_2} = \rho$ we can test the hypothesis that $\mu_{x_1} = \mu_{y_1}$ and $\mu_{x_2} = \mu_{y_2}$ simultaneously against the alternative that $\mu_{x_1} \neq \mu_{y_1}$ and $\mu_{x_2} \neq \mu_{y_2}$. Let the the differences between the sample means be denoted by $d_i = \bar{x}_i - \bar{y}_i$, $i = 1, 2$. These differences, under the null hypothesis, are normally distributed with mean values zero and variances

$$\sigma_i^2\left(\frac{1}{k_1} + \frac{1}{k_2}\right), \qquad i = 1, 2.$$

and correlation coefficient ρ. σ_i^2 and ρ can be estimated by

$$s_i^2 = \frac{(k_1-1)s_{x_i}^2 + (k_2-1)s_{y_i}^2}{k_1 + k_2 - 2} \quad \text{and}$$

$$r = \frac{(k_1-1)s_{x_1x_2} + (k_2-1)s_{y_1y_2}}{k_1 + k_2 - 2} \Big/ s_1 s_2$$

where

$$s_{x_1x_2} = \frac{1}{k_1-1}\sum_{j=1}^{k_1}(x_{1j} - \bar{x}_1)(x_{2j} - \bar{x}_2)$$

and $s_{y_1y_2}$ is defined similarly. To test the hypothesis that d_1 and d_2 come from populations with means zero against the alternative that they do not come from populations with means zero, we form the expression

$$T^2 = \frac{1}{1-r^2}\left[\frac{d_1^2}{s_{d_1}^2} - 2r\,\frac{d_1 d_2}{s_{d_1}s_{d_2}} + \frac{d_2^2}{s_{d_2}^2}\right] \qquad 6.7$$

In 6.7 $s_{d_1}^2$ and $s_{d_2}^2$ are estimates of the variances of d_1 and d_2 and are given by

$$s_i^2\left(\frac{1}{k_1} + \frac{1}{k_2}\right), \qquad i = 1, 2$$

T^2 is distributed as

$$\frac{2\,(k_1 + k_2 - 2)}{(k_1 + k_2 - 3)}F_{2,\; k_1+k_2-3}$$

under suitable normality assumptions. Again, an $\alpha\%$ test is made in the usual way making use of the appropriate F tables.

6.3 Relation between Trace B and Hotelling's T^2

Returning now to trace B as defined in Chapter 3 it is a small exercise in algebraic manipulation to relate trace B and T^2. Recall that (x_i, y_i), $1 \leq i \leq k$ are a set of difference vectors. One transformation taking (x_i, y_i) into (a_i, b_i) such that

$\sum_{i=1}^{k} a_i^2 = \sum_{i=1}^{k} b_i^2 = 1$ and $\sum_{i=1}^{k} a_i b_i = 0$ is

$$a_i = \frac{x_i}{\sqrt{\sum_{i=1}^{k} x_i^2}} \quad \text{and} \quad b_i = \frac{y_i - \dfrac{\sum_{i=1}^{k} x_i y_i}{\sum_{i=1}^{k} x_i^2}\, x_i}{\sqrt{\sum_{j=1}^{k}\left(y_j - \dfrac{\sum_{i=1}^{k} x_i y_i}{\sum_{i=1}^{k} x_i^2}\, x_j\right)^2}} \qquad 6.8$$

Substituting 6.8 in the formula for trace B and denoting $\sum_{i=1}^{k}$ simply by Σ we obtain

$$\text{trace } B = \frac{1}{k}[(\Sigma a_i)^2 + (\Sigma b_i)^2]$$

$$= \frac{1}{k}\left[\frac{(\Sigma x_i)^2}{\Sigma x_i^2} + \frac{\left(\Sigma y_i - \dfrac{\Sigma x_i y_i}{\Sigma x_i^2}\, \Sigma x_i\right)^2}{\Sigma\left(y_i - \dfrac{\Sigma x_i y_i}{\Sigma x_i^2}\, x_i\right)^2}\right]$$

$$= \frac{1}{k}\; \frac{(\Sigma x_i)^2 \Sigma y_i^2 \Sigma x_i^2 - (\Sigma x_i)^2 (\Sigma x_i y_i)^2 + (\Sigma x_i^2)^2 (\Sigma y_i)^2 - 2\Sigma x_i y_i \Sigma x_i \Sigma y_i \Sigma x_i^2 + (\Sigma x_i y_i)^2 (\Sigma x_i)^2}{\Sigma x_i^2 [\Sigma x_i^2 \Sigma y_i^2 - (\Sigma x_i y_i)^2]}$$

$$= \frac{1}{k}\left[\frac{(\Sigma x_i)^2 \Sigma y_i^2 + \Sigma x_i^2 (\Sigma y_i)^2 - 2\,\Sigma x_i y_i \Sigma x_i \Sigma y_i}{\Sigma x_i^2 \Sigma y_i^2 - (\Sigma x_i y_i)^2}\right]$$

Let

$$\Sigma x_i = k\bar{x} \qquad \Sigma y_i = k\bar{y}$$

$$\Sigma x_i^2 = (k-1)s_x^2 + k\bar{x}^2 \qquad \Sigma y_i^2 = (k-1)s_y^2 + k\bar{y}^2$$

$$\Sigma x_i y_i = (k-1)s_{xy} + k\bar{x}\bar{y} \qquad r = s_{xy}/s_x s_y$$

Then

$$\text{trace } B = \frac{k^2\bar{x}^2[(k-1)s_y^2 + k\bar{y}^2] + k^2\bar{y}^2[(k-1)s_x^2 + k\bar{x}^2] - 2[(k-1)s_{xy} + k\bar{x}\bar{y}]k^2\bar{x}\bar{y}}{k\{[(k-1)s_x^2 + k\bar{x}^2][(k-1)s_y^2 + k\bar{y}^2] - [(k-1)s_{xy} + k\bar{x}\bar{y}]^2\}}$$

$$= \frac{k^2\bar{x}^2 s_y^2 + k^2\bar{y}^2 s_x^2 - 2k^2\bar{x}\bar{y}s_{xy}}{k[(k-1)(s_x^2 s_y^2 - s_{xy}^2) + k\bar{x}^2 s_y^2 + k\bar{y}^2 s_x^2 - 2k\bar{x}\bar{y}s_{xy}]}$$

$$= \frac{1}{(k-1)\dfrac{1-r^2}{k\dfrac{\bar{x}^2}{s_x^2} - 2kr\dfrac{\bar{x}\bar{y}}{s_x s_y} + k\dfrac{\bar{y}^2}{s_y^2}} + 1}$$

From formula 6.5 we see that

$$T^2 = \frac{1}{1-r^2}\left[k\frac{\bar{x}^2}{s_x^2} - 2kr\frac{\bar{x}\bar{y}}{s_x s_y} + k\frac{\bar{y}^2}{s_y^2}\right]$$

Thus

$$\text{trace } B = \frac{1}{(k-1)\frac{1}{T^2} + 1} = \frac{T^2}{(k-1) + T^2}$$

6.4 Distribution of trace B Under Normality Assumptions

The distribution of T^2 under suitable normality assumptions for p dimensions and k replications is given by

$$f(T^2)dT^2 = \frac{1}{B(\frac{k-p}{2}, \frac{p}{2})} \left(\frac{T^2}{k-1} \right)^{\frac{p}{2}-1} \left(1 + \frac{T^2}{k-1} \right)^{-\frac{k}{2}} d\left(\frac{T^2}{k-1} \right)$$

Let trace B when the underlying distribution is normal be denoted by W_N. Since

$$W_N = \frac{T^2}{(k-1) + T^2}$$

for the case of two dimensions, one arrives at

$$f(W_N)dW_N = \frac{1}{B(\frac{k-2}{2}, 1)} \left(\frac{W_N}{1-W_N} \right)^{\frac{2}{2}-1} \left(\frac{1}{1-W_N} \right)^{-\frac{k}{2}+2} dW_N$$

$$= \frac{1}{B(\frac{k-2}{2}, 1)} \left(1 - W_N \right)^{\frac{k}{2}-2} dW_N \qquad 6.9$$

for the distribution of W_N.

6.5 An Empirical Sampling Experiment

For the purpose of investigating empirically the effect of non-normality on Hotelling's T^2, a sampling experiment using random numbers and random normal deviates was performed. Forty-eight samples of size 8 were drawn from both a two-dimensional uniform distribution on $(-\frac{1}{2}, \frac{1}{2})$ and from a bivariate normal distribution with zero mean, unit variance and zero covariance.

The sample size 8 was chosen for several reasons. It was a manageable size for the necessary computations (involving 2^7 randomizations for each sample) which could be conveniently programmed for the electronic computer available (Burroughs E101). Also, it seems not unlikely that samples of size 8 often occur in actual problems. Furthermore, for a sample as small as size 8 one becomes somewhat concerned about making the necessary normality assumptions for a T^2 test.

Forty-eight samples were chosen because it seemed that a reasonably large number of samples would be required to make definite conclusions about significance points; but at the same time forty-eight was small enough to keep the amount of computing within reason. Samples from a rectangular distribution were chosen because of the convenience in using random numbers and the desire to investigate a distribution considerably different from the normal distribution. Samples from a normal distribution were chosen to both verify the results arrived at by direct mathematical means and at the same time to clearly show the contrast between frequency distributions of significance points obtained from samples from a non-normal distribution and those obtained from samples from a normal distribution.

An individual pair of elements from any sample of size 8 was considered as being the observed difference of the measurements within a block from an experiment consisting of eight blocks, each block containing the results of the quantitative measurement of two responses from two different treatments. These differences (let them be denoted by x_i, y_i) are then rescaled in accordance with the transformation given by formula 6.8 with $k = 8$. For each sample

$$8 \text{ trace } B = (\sum_{i=1}^{8} a_i)^2 + (\sum_{i=1}^{8} b_i)^2$$

was computed as were all the 127 other traces obtained by not changing and changing the sign of each pair of (a_1, b_1). This gives rise to a set of 128 quantities ranging from 0 to 8 with an average value of 2.

With the 128 values of 8 trace B for each sample, one can pick out the largest values and then compare the empirical distribution with that expected under normality assumptions. A frequency distribution of the number of values of $W = 8$ trace B which exceed the $\alpha\%$ point of W_α can be constructed. The frequency distributions for both the samples from a normal distribution and from a rectangular distribution for the 5.0, 2.5, 1.0 and 0.5% points are given in Table 6.1. Table 6.2 gives the actual percentage points of the empirical distributions obtained by computing the mean of the distributions in Table 6.1 and converting to percentages. Also 95% confidence intervals are given in Table 6.2. It follows from Table 6.2 that an approximate rule for computing the actual significance level when a test is carried out at the $\alpha\%$ $(\alpha > 0)$ level on data which deviate from normality in the same way that the uniform distribution deviates from normality is simply

$$\text{actual significance level} = 1.1\,\alpha + 0.5$$

This formula gives actual significance levels of 1.05, 1.60, 3.25 and 6.00 in place of the observed 0.96, 1.68, 3.24 and 6.02.

6.6 Adjustment of Parameters in Distribution of Trace B

In the foregoing example, one is left with the possibility of altering the parameters in the distribution of trace B so as to make a more nearly exact significance test. This can be done in

Table 6.1

Frequency distribution of upper α% tail based on normal theory for 48 samples from the rectangular distribution and from the normal distribution

No. exceeding α% point	Normal Samples				Rectangular Samples			
	5%	2.5%	1.0%	0.5%	5%	2.5%	1.0%	0.5%
0		1	14	22		1	2	8
1		4	13	17	1	1	13	24
2	3	11	12	2	0	2	16	13
3	1	7	8		0	10	10	3
4	4	9	1		1	14	7	
5	7	12			4	12		
6	9	4			4	8		
7	6				12			
8	7				9			
9	6				9			
10	3				4			
11	2				4			

Table 6.2

Actual significance level of α% test for samples from normal and rectangular distribution

	Normal Samples				Rectangular Samples			
Nominal %	5.0	2.5	1.0	0.5	5.0	2.5	1.0	0.5
Actual %	5.18	2.72	1.05	0.60	6.02	3.24	1.68	0.96
95% Confidence Interval	4.66 to 5.70	2.36 to 3.08	0.79 to 1.31	0.40 to 0.80	5.58 to 6.46	2.94 to 3.54	1.42 to 1.94	0.78 to 1.14

a number of ways by making certain assumptions on the form of the beta distribution to be fitted. The most general model that one can fit makes use of the first four cumulants of trace B. This requires the fitting of the four parameters in the distribution given by

$$f(W)dW = C\left(\frac{W}{a} - 1\right)^{\mu-1}\left(1 - \frac{W}{b}\right)^{\nu-1} \qquad a < W < b \qquad 6.10$$

where C is a constant such that

$$\int_a^b f(W)dW = 1$$

Since we know that trace B is a random variable between 0 and 1 it seems that a convenient thing to do is to fit μ and ν by the first two cumulants. The problem then becomes one of finding the α% points of a distribution of the form

$$f(W)dW = \frac{1}{B(\mu, \nu)} W^{\mu-1}(1 - W)^{\nu-1}\, dW \qquad 6.11$$

In the case of samples from a normal distribution $\mu = 1$ and $\nu = 3$, so one is then required to find the α % points of 6.11 in the neighbourhood of $\mu = 1$ and $\nu = 3$. Since the Tables of the Incomplete Beta Functions give only values of μ and ν by steps of 0.5 and since interpolation for such small values of μ and ν is very unsatisfactory, one must approximate these percentage points. An approximate method of finding the necessary percentage points is given in the Appendix together with a chart (Figure A.1) showing the upper 5.0, 2.5, 1.0 and 0.5% curves of δW over a range of $\sum_{p=1}^{8}(a_p^2 + b_p^2)^2$.

Note that because the mean of the 128 permutations of trace B is 1/4 for every sample, the ratio of μ to ν is always 1/3. By making use of Figure A.1, one can find adjusted percentage points to apply to the randomizations of the samples from the rectangular distribution. The frequency distributions are given in Table 6.3.

Table 6.3

Frequency distributions of upper α% tail based on adjusting the degrees of freedom in the unconditional distribution of W by means of the first two cumulants of the conditional distribution of W for samples from the rectangular distribution.

No. exceeding α% point	Rectangular Samples			
	5%	2.5%	1.0%	0.5%
0		1	4	14
1		1	21	29
2		4	18	5
3		18	5	
4	2	19		
5	8	2		
6	12	3		
7	15			
8	7			
9	3			
10	1			

Table 6.4

Actual significance level of α% test for samples from the rectangular distribution, adjusting μ and ν by the first two cumulants

Nominal %	Actual %	95% Confidence Interval
5.0	5.18	4.88 to 5.48
2.5	2.72	2.46 to 2.98
1.0	1.17	0.99 to 1.35
0.5	0.63	0.51 to 0.75

Figure 6.1

Actual 5% values of permutation distributions of 8 trace B for 48 samples of size 8 from a rectangular distribution and from a normal distribution

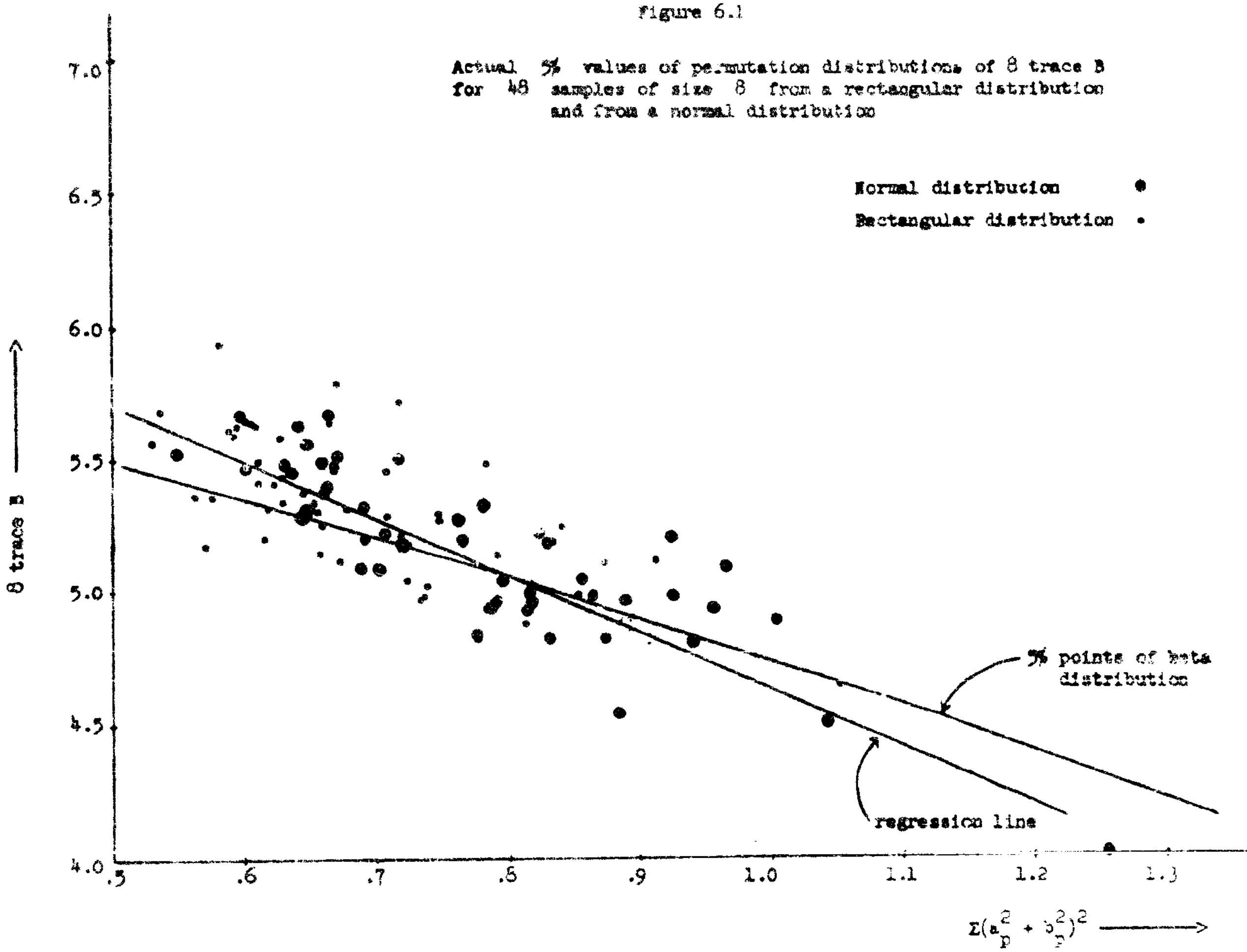

Table 6.4 shows the significance levels after making the adjustment and one can see that this method of adjustment works quite well in this instance. Figure 6.1 shows the individual points obtained by selecting a value from the permutation distribution in each of the 96 samples which is such that 5% of the values of 8 trace B are larger than the value selected. The fact that the points obtained from the samples from the rectangular distribution when plotted against $\sum_{p=1}^{8}(a_p^2 + b_p^2)^2$ (a function of the variance of trace B only) have a similar pattern to those obtained from the samples from the normal distribution, is further justification for fitting the upper tail of these empirical permutation distributions obtained from the samples from the rectangular distributions by the first and second cumulants only. In fact the regression lines fitted through $\sum_{p=1}^{8}(a_p^2 + b_p^2)^2 = 0.8$ and 8 trace $B = 5.05$, the point corresponding to $\mu = 1$ and $\nu = 3$, are almost identical for the two empirical permutation distributions.

A plot of the deviations of the actual 5% points from those based on the fitted beta distribution against the sixth degree invariant $(\sum_{p=1}^{k}(a_p^2 + b_p^2)^3)$ shows that little or nothing would be gained by fitting the beta distribution by the third cumulant also.

Other ways of adjusting the beta distribution to fit the upper tail of the empirical distributions were also tried. However, these methods proved to be considerably inferior to the one just outlined. They are presented here to indicate ways of improving the significance level of the test when one does not have the time or inclination to approximate the percentage points of the incomplete beta function.

ERRATA

PAGE	LINE	IN PLACE OF	READ
[illegible]	6	x_{ij} [illegible]	x_{ij} [illegible]
[illegible]	[illegible]*	[illegible]ments	treatments
[illegible]	5	[illegible]	[illegible]
6	6	$a_{1i}\ a_{1i}a_{12}\ \ldots\ a_{1i}a_{1m}$	$a_{1i}^2\ a_{1i}\ a_{12}\ \ldots\ a_{1i}a_{1m}$
[illegible]	[illegible]b	ave (rot) $\sum_{p,q=1}^{n} a_p^2 a_q^2$	ave (rot) $\sum_{p,q=1}^{k} a_p^2 a_q^2$
47	Table 5.1	(3) [illegible] (4)	(?) [illegible] (5)
[illegible]	4c	$(x_i - \mu_i)^2$	$(\bar{x}_i - \mu_i)^2$
6[illegible]	5b	μ_{y1}, μ_{y2}	μ_{y1}, μ_{y2}
51	7b	let the	let
76	8b	oppostie	opposite
80	Table A 1	(trace B)	3(trace B)
82	Table A 2	B(trace B)	8(trace B)
84	Table A 3	$F(a_p b_q \ldots)^4$	$F^*(a_p b_q - b_p a_q)^4$

* As counted from the bottom.

6.7 Further Methods for Adjusting Significance Levels

The criterion

$$W = \text{trace } B = \frac{1}{k}\left[\left(\sum_{1}^{k} a_p\right)^2 + \left(\sum_{1}^{k} b_p\right)^2\right]$$

might be considered as being distributed as

$$f(W)dW = c\left(1 - \frac{W}{a}\right)^{\nu-1} dW \qquad 0 < W < a \qquad 6.12$$

There are many reasons for rejecting this distribution from consideration. The main reason for rejecting is probably because we know that the upper limit of the range of trace B is 1 but this distribution given by 6.12 has an upper limit of a. Thus one might arrive at a $\alpha\%$ significance point for trace B in a particular case which is outside of the range $(0, 1)$. However, this distribution does allow one to alter the significance points with a very small amount of computation. To fit this distribution one needs to find a, ν and c such that

$$\int_0^a c\left(1 - \frac{W}{a}\right)^{\nu-1} dW = 1$$

$$\int_0^a c\, W\left(1 - \frac{W}{a}\right)^{\nu-1} dW = \frac{1}{4}$$

$$\int_0^a c\, W^2\left(1 - \frac{W}{a}\right)^{\nu-1} dW = \frac{1}{32}\left[2 - \sum_{1}^{k}(a_p^2 + b_p^2)^2\right] + \frac{1}{16}$$

These equations when solved give

$$\nu = \frac{4}{\sum_{1}^{k}(a_p^2 + b_p^2)^2} - 2$$

and $$a = \frac{1}{\sum_{1}^{k}(a_p^2 + b_p^2)^2} - \frac{1}{4}$$

These formulas applied to the 48 samples from the uniform distribution give rise to a set of ν's and a's each of which are used to obtain a value of W_α by the formula

$$\left(1 - \frac{W_\alpha}{a}\right)^{\nu} = \alpha$$

or $$W_\alpha = a(1 - \alpha^{1/\nu})$$

One then records the number of values in the randomization distribution of trace B which exceeds this W_α. The frequency distributions are given in Table 6.5 and associated adjusted percentage points and confidence intervals in Table 6.6. It is interesting to note that the nominal 5% and 2.5% significance levels for the case of the samples from the uniform distribution give rise to actual observed significance levels of 5.17% and 2.46% as opposed to the 6.02% and 3.24% arrived at without any adjustment. However, the 1% and .5% levels do not have this rather good agreement.

An alternative method is to fix the upper end of the distribution at 1 and allow the lower end to wander. This again requires the use of only the first two cumulants of B and without further details, the results are given in Tables 6.7 and 6.8.

Neither of these two approaches gives rise to very satisfactory results. However the first of the two does give reasonable results at the 5% and 2.5% levels at the cost of bad results at the 1% and 0.5% levels. If one were interested in a test at the 5% level and wished to make an improvement in the accuracy of the actual

Table 6.5

Frequency distributions of upper α% tail based on allowing the lower end of the distribution of trace B to be fixed at zero and adjusting the upper end and the parameter ν for samples from the rectangular distribution

No. exceeding α% point	Rectangular Samples 5%	2.5%	1.0%	0.5%
0			14	38
1		2	29	9
2		13	4	1
3		14	1	
4	2	14		
5	9	5		
6	11			
7	14			
8	8			
9	3			
10	1			

Table 6.6

Actual significance level of α% test for samples from the rectangular distribution, adjusting ν and upper end of distribution of trace B

Nominal %	Actual %	95% Confidence Interval
5	5.17	4.87 to 5.47
2.5	2.46	2.22 to 2.70
1.0	0.65	0.51 to 0.79
0.5	0.10	0.00 to 0.20

Table 6.7

Frequency distributions of upper $\alpha\%$ tail based on allowing the upper end of the distribution of trace B to be fixed at one and adjusting the lower end and the parameter ν for samples from the rectangular distribution

No. exceeding $\alpha\%$ point	Rectangular Samples 5%	2.5%	1.0%	0.5%
0		1	2	12
1		1	20	27
2		3	17	8
3		14	9	1
4	2	24		
5	8	1		
6	8	4		
7	16			
8	7			
9	7			

Table 6.8

Actual significance level of $\alpha\%$ test for samples from the rectangular distribution, adjusting ν and lower end of distribution of trace B

Nominal %	Actual %	95% Confidence Interval
5.0	5.32	5.01 to 5.63
2.5	2.83	2.57 to 3.09
1.0	1.32	1.14 to 1.50
0.5	0.75	0.59 to 0.91

significance level, then this method does offer a fairly simple and probably reliable answer. Little can be said in favor of using the second of the two methods except that the rather poor improvement in the actual significance level is uniform for all four levels investigated.

Any other method of fitting using the beta distribution requires the use of more extensive tables than the ones now available for obtaining the percentage points of the beta distributions. If these tables were available over a fine grid for small numbers of degrees of freedom, then one could test the adequacy of fitting by only the first two cumulants versus the use of all the four cumulants which have been theoretically determined and can be computed rather easily. Tables of the necessary computations for obtaining the first four cumulants of trace B for all 50 samples are given in the Appendix (Tables A.3 and A.4).

6.8 Suggestions for Further Investigations

Although certain conclusions seem evident from the limited empirical study done here, it seems that in order to fully investigate the effect of non-normality on Hotelling's T^2 one needs to investigate other distributions than the rectangular. A distribution which would seem to be worth considering is the double exponential. This would give some information on a distribution which differs from the normal distribution in the opposite way to that in which the rectangular differs.

Clearly, there is no end to the number of sample sizes that one might use in the permutation approach. However, results based on samples of size 8 only are certainly inadequate in drawing conclusions about samples of sizes other than 8 without first investigating what happens for several different sample sizes. To investigate samples much larger than size 8 one needs at least a medium size electronic computer.

The problem of investigating Hotelling's T^2 where there are more than two responses being measured was not treated here but a small amount of modification in the methods used should produce the necessary results. Computationally, more than two responses should create no new problems.

APPENDIX

A.1 Method of Approximating the Percentage Points of the Beta Distribution Over a limited range

We require W_α for $\alpha = 0.05,\ 0.025,\ 0.01,\ 0.005$ where

$$\frac{1}{B(\mu,\ \nu)} \int_{W_\alpha}^{1} W^{\mu-1} (1 - W)^{\nu-1}\, dW = \alpha \qquad \text{A.1}$$

over a range of $(\mu,\ \nu)$ such that $\nu = 3\mu$ and μ lies between 0.5 and 2.0. If we let $1 - W = t$ then A.1 becomes

$$\frac{1}{B(\mu,\ \nu)} \int_{0}^{1-W_\alpha} (1 - t)^{\mu-1}\, t^{\nu-1}\, dt = \alpha$$

Expanding $(1 - t)^{\mu-1}$ by a Taylor's series we obtain

$$\frac{1}{B(\mu,\ \nu)} \int_{0}^{1-W_\alpha} [1 - (\mu-1)t + (\mu-1)(\tfrac{\mu-2}{2})t^2 - (\mu-1)(\tfrac{\mu-2}{2})(\tfrac{\mu-3}{3})t^3]t^{\nu-1}\, dt \approx \alpha \qquad \text{A.2}$$

Replacing $1 - W_\alpha$ by t_α and integrating term by term A.2 is replaced by

$$\frac{t_\alpha^{\nu}}{\nu} - (\mu-1)\frac{t_\alpha^{\nu+1}}{\nu+1} + (\mu-1)(\tfrac{\mu-2}{2})\frac{t_\alpha^{\nu+2}}{\nu+2} - (\mu-1)(\tfrac{\mu-2}{2})(\tfrac{\mu-3}{3})\frac{t_\alpha^{\nu+3}}{\nu+3} \approx \alpha B(\mu,\ \nu)$$

Applying Newton's method for finding the root of a polynomial, t_α can be quickly found over the necessary range of $\nu = 3\mu$. The accuracy of this approximation was checked at each end of the range of μ required (0.5 and 2.0) against the values obtained by linear interpolation in the Tables of the Incomplete Beta Function. The approximation agreed to at least four significant digits for all values of α. A chart, constructed from the computed points for $\mu = .7,\ .8,\ .9,\ 1.1,\ 1.2,\ 1.3,\ 1.4,\ 1.5$ and 2.0 is given as Figure A.1.

A.2 Comment on Tables A.1, A.2

In some of the samples, more than the ten largest values of the permutation distribution are given. These extra values are given because these samples possessed more than ten permutations exceeding the 5% normal theory significance point.

2

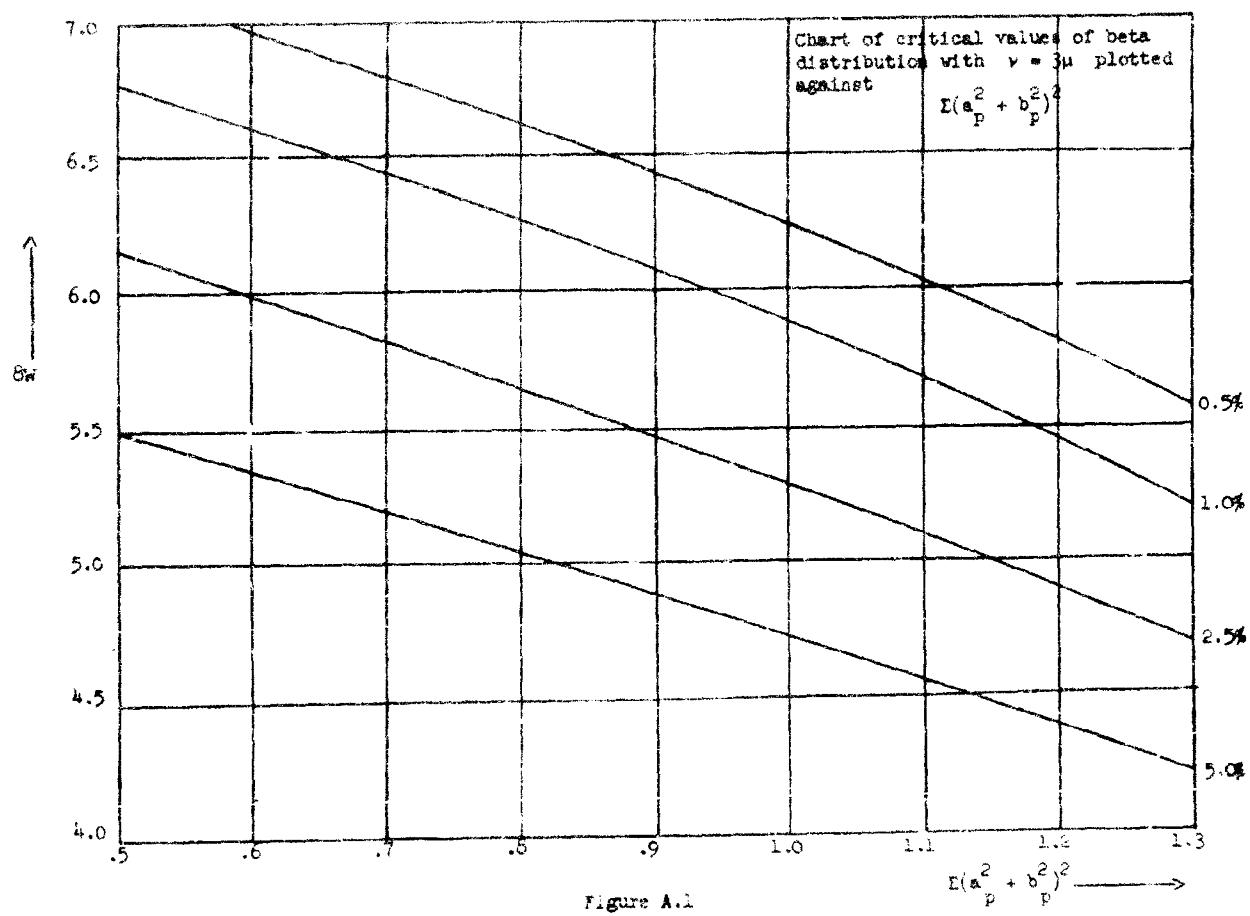
Chart of critical values of beta distribution with $\nu = 3\mu$ plotted against $\Sigma(a_p^2 + b_p^2)^2$
7.0
6.5
6.0
5.5
5.0
4.5
4.0
8_W
.5
.6
.7
.8
.9
1.0
1.1
1.2
1.3
$\Sigma(a_p^2 + b_p^2)^2$
0.5%
1.0%
2.5%
5.0%

Figure A.1

Table A.1

The Largest Values of the Permutation Distribution of (trace B) for 48 Samples of size 8 from the Uniform Distribution

SAMPLE NO:

1	2	3	4	5	6	7	8
7.240	5.115	6.503	6.961	6.682	7.180	7.286	7.579
6.763	4.998	6.067	6.868	6.525	5.713	6.761	7.224
6.649	5.990	5.414	6.664	6.525	5.589	6.315	6.620
6.411	4.941	5.360	6.301	6.473	5.524	6.032	5.888
6.249	4.836	5.208	5.710	5.789	5.513	5.670	5.437
6.153	4.672	5.158	5.416	5.363	5.352	5.362	5.410
5.432	4.620	5.112	5.395	5.348	5.152	5.289	5.288
5.357	4.617	5.091	4.901	5.259	4.856	5.005	5.274
5.041	4.488	5.067	4.733	5.222	4.852	4.979	4.992
5.031	4.475	5.036	4.638	5.214	4.634	4.974	4.987

SAMPLE NO:

9	10	11	12	13	14	15	16
7.555	5.764	6.305	6.960	6.994	7.277	7.280	7.205
6.130	5.589	5.768	6.437	6.618	6.944	6.955	6.677
6.051	5.535	5.685	5.708	6.542	6.073	6.480	6.044
5.730	5.518	5.657	5.502	6.074	6.071	5.658	5.866
5.503	5.348	5.631	5.419	5.719	5.661	5.360	5.622
5.345	5.321	4.901	5.262	5.670	5.642	5.213	5.385
5.331	4.829	4.855	5.169	5.602	5.618	5.168	5.231
5.153	4.766	4.773	5.106	5.491	5.538	5.130	5.219
4.832	4.675	4.641	5.091	5.375	5.006	5.108	5.181
4.815	4.385	4.584	5.011	5.095	5.003	5.040	5.059

SAMPLE NO:

17	18	19	20	21	22	23	24
6.563	6.993	6.816	6.842	6.693	7.620	7.172	6.659
6.185	6.802	6.650	6.365	6.321	6.701	5.792	6.242
6.053	6.435	6.021	6.275	6.274	6.540	5.719	6.172
5.997	6.384	5.920	5.864	6.060	6.202	5.339	5.584
5.551	5.848	5.527	5.513	5.475	6.017	5.104	5.271
5.422	5.598	5.392	5.302	5.473	5.712	5.042	5.198
4.971	5.589	5.158	5.198	5.408	5.544	4.885	5.185
4.928	5.469	5.087	5.134	5.266	5.013	4.805	5.151
4.650	5.440	4.702	5.104	5.062	4.614	4.626	4.782
4.396	4.903	4.668	5.050	5.019	4.598	4.48[illegible]	4.396

Table A.1 (contd.)

SAMPLE NO:

25	26	27	28	29	30	31	32
6.725	6.798	7.444	7.166	6.607	7.296	6.809	6.681
6.364	6.092	6.450	6.644	6.221	6.962	6.242	6.628
5.982	5.755	6.341	6.432	5.914	5.520	6.130	6.187
5.756	5.540	6.018	5.519	5.706	6.517	5.725	6.150
5.138	5.527	5.868	5.386	5.704	5.877	5.425	5.754
5.039	5.368	5.815	5.087	5.340	5.749	5.424	5.696
5.034	5.168	5.579	4.953	5.296	5.587	5.062	5.505
5.023	5.045	4.904	4.879	5.145	5.395	4.901	5.433
4.840	4.828	4.583	4.871	4.944	5.365	4.884	5.209
4.786	4.602	4.573	4.799	4.744	5.209	4.876	5.201
					5.184		5.171

SAMPLE NO:

33	34	35	36	37	38	39	40
7.118	7.807	6.486	6.682	7.524	7.251	7.262	7.655
5.985	6.346	6.375	5.925	6.801	6.627	6.598	6.322
5.978	6.067	6.328	5.790	6.282	6.153	5.669	5.960
5.965	5.879	6.095	5.719	6.178	5.937	5.470	5.926
5.929	5.737	6.078	5.635	5.885	5.690	5.247	5.732
5.859	5.449	5.964	5.631	5.189	5.493	5.179	5.728
5.306	5.400	5.509	5.236	5.142	5.282	5.011	5.486
4.809	5.117	5.287	4.546	5.015	5.235	4.969	5.389
4.625	4.887	5.051	4.531	4.946	5.189	4.967	5.369
4.348	4.731	4.930	4.462	4.772	4.787	4.720	5.206
							5.153

SAMPLE NO:

41	42	43	44	45	46	47	48
7.553	7.491	7.142	6.788	7.176	6.751	6.448	7.179
7.018	6.355	6.909	6.435	7.002	6.107	6.293	6.351
6.972	5.662	6.569	5.935	6.416	6.064	6.165	6.151
6.354	5.630	6.294	5.704	6.227	5.930	5.713	5.615
5.694	5.616	5.461	5.616	5.670	5.039	5.386	5.286
5.652	5.580	5.202	5.384	5.511	5.008	5.162	5.047
5.444	5.574	5.061	5.344	5.461	4.948	5.056	4.842
5.411	5.538	4.805	5.263	5.317	4.720	4.919	4.668
5.233	5.380	4.738	5.247	5.268	4.645	4.887	4.652
4.900	5.346	4.698	5.182	4.615	4.618	4.884	4.425
	5.061						

Table A.2

The Largest Values of the Permutation Distribution of 8(tract B) for 48 Samples of size 8 from the Normal Distribution

SAMPLE NO:

1	2	3	4	5	6	7	8
6.165	5.885	7.323	5.983	7.034	6.939	6.827	7.245
5.727	5.172	6.811	5.729	6.559	6.858	6.731	6.144
5.179	4.779	6.049	5.653	6.363	6.851	6.530	5.840
5.115	4.432	5.780	5.487	6.072	6.444	5.894	5.730
5.043	4.148	5.763	5.063	6.036	6.148	5.770	5.685
5.001	4.091	5.515	4.892	5.513	5.527	5.743	5.549
4.737	3.865	5.478	4.785	5.237	5.423	5.458	5.368
4.535	3.857	5.314	4.650	5.231	5.129	5.208	5.175
4.357	3.855	5.287	4.623	5.157	4.675	5.072	5.166
4.297	3.819	4.971	4.621	4.889	4.585	5.034	5.102

SAMPLE NO:

9	10	11	12	13	14	15	16
7.134	6.783	5.726	6.147	6.520	5.889	6.693	5.808
6.678	6.713	5.583	5.738	6.456	5.568	5.962	5.708
6.043	6.287	5.470	5.730	6.373	5.535	5.851	5.628
5.827	5.318	5.290	5.340	5.978	5.343	5.756	5.410
5.794	5.076	5.182	5.116	5.958	5.311	5.739	5.241
5.509	4.901	5.114	5.016	5.565	5.268	5.588	4.872
5.399	4.733	4.837	5.013	5.545	5.157	4.976	4.835
5.305	4.781	4.704	4.982	5.289	5.081	4.967	4.732
5.128	4.780	4.676	4.862	4.955	5.068	4.876	4.667
5.055	4.309	4.558	4.781	4.724	4.776	4.763	4.581

SAMPLE NO:

17	18	19	20	21	22	23	24
6.140	7.039	6.370	7.673	6.776	6.728	7.021	6.047
5.728	6.913	6.130	6.631	6.096	6.359	6.605	5.727
5.631	6.808	5.816	6.215	5.777	6.051	6.228	5.689
5.552	6.267	5.566	6.133	5.698	5.858	5.972	5.190
5.486	6.044	5.499	5.402	5.553	5.825	5.268	5.093
5.416	5.544	5.375	5.115	5.114	5.697	5.225	5.071
4.927	5.499	4.965	5.043	5.105	5.604	5.122	4.865
4.901	5.448	4.928	4.979	5.060	5.486	5.049	4.703
4.845	5.257	4.757	4.904	4.986	5.103	4.848	4.439
4.829	5.231	4.756	4.592	4.844	4.857	4.572	4.306
	5.096						

Table A.2 (contd.)

SAMPLE NO:

25	26	27	28	29	30	31	32
6.633	7.699	6.436	6.459	6.352	7.049	6.116	6.528
6.387	6.481	6.233	6.014	5.680	6.350	5.800	6.452
5.212	6.267	5.920	5.618	5.472	6.227	5.275	6.206
5.108	5.758	5.595	5.324	5.428	6.191	5.214	6.052
5.021	5.316	5.473	5.235	5.386	6.048	4.952	5.513
4.988	5.113	4.833	5.036	5.316	5.368	4.942	5.229
4.978	4.966	4.822	4.896	5.224	5.230	4.930	5.172
4.732	4.783	4.690	4.722	5.194	5.182	4.696	5.138
4.703	4.583	4.607	4.696	5.007	4.941	4.418	4.960
4.630	4.439	4.481	4.598	4.963	4.795	4.373	4.919

SAMPLE NO:

33	34	35	36	37	38	39	40
6.161	7.139	7.127	7.234	6.936	7.014	5.528	6.639
5.465	6.775	6.671	6.463	6.134	6.618	5.313	6.544
5.044	6.618	5.980	6.329	6.079	6.611	4.958	6.245
4.949	6.030	5.772	6.173	5.977	5.623	4.813	4.982
4.948	5.911	5.770	6.046	5.728	5.557	4.749	4.965
4.883	5.756	5.741	5.588	5.337	5.396	4.522	4.937
4.766	5.506	5.198	5.404	5.252	5.219	4.508	4.914
4.520	5.329	5.131	5.005	5.007	5.012	4.498	4.875
4.429	5.122	4.987	4.764	4.934	4.748	4.482	4.580
4.230	5.098	4.785	4.722	4.956	4.611	4.475	4.485

SAMPLE NO:

41	42	43	44	45	46	47	48
7.258	6.631	6.838	6.890	6.585	6.028	7.175	6.267
6.041	6.074	6.737	5.956	6.219	5.885	6.127	5.949
5.827	5.520	5.510	5.903	6.040	5.848	5.745	5.467
5.740	5.270	5.432	5.773	5.949	5.803	5.702	5.284
5.624	5.245	5.125	5.542	5.749	5.132	5.489	4.675
5.292	5.091	5.084	5.410	5.240	4.959	5.486	4.599
5.122	5.007	4.838	5.302	5.131	4.931	5.438	4.541
5.011	4.747	4.816	5.249	5.002	4.870	5.342	4.489
4.948	4.573	4.806	5.055	4.956	4.846	5.182	4.303
4.927	4.453	4.703	4.928	4.736	4.654	5.073	4.242

Table A.3

Values of the Six Invariants for Computing the First Four Cumulants of trace B for 45 Samples of size 8 from the Uniform Distribution.

Sample No	$\Sigma(a_p^2+b_p^2)^2$	$\Sigma(a_p^2+b_p^2)^3$	$\Sigma(a_p^2+b_p^2)^4$	$[\Sigma(a_p^2+b_p^2)^2]^2$	$\Sigma''(a_pa_q+b_pb_q)^4$	$\Sigma''(a_pb_q+b_pa_q)^4$
1	.5815	.2000	.0795	.3382	.0627	.1251
2	1.0525	.6624	.4409	1.1077	.0680	.4912
3	.7917	.3853	.2120	.6268	.0705	.2465
4	.6057	.2015	.0711	.3669	.0753	.1346
5	.5752	.1804	.0501	.3308	.0711	.1290
6	.7481	.3367	.1704	.5597	.0813	.2162
7	.6282	.2303	.0941	.3947	.0663	.1433
8	.5608	.1706	.0546	.3145	.0893	.1432
9	.6521	.2413	.0951	.4253	.0935	.1797
10	.9124	.5164	.3245	.8326	.0521	.3117
11	.8126	.3907	.2010	.6603	.0689	.2543
12	.7157	.3137	.1561	.5122	.0675	.1959
13	.6046	.2015	.0703	.3655	.0808	.1489
14	.6067	.2054	.0741	.3681	.0864	.1514
15	.6124	.2088	.0756	.3751	.0908	.1637
16	.6165	.2074	.0728	.3800	.0782	.1455
17	.8391	.4566	.2793	.7041	.0732	.2948
18	.5892	.1930	.0674	.3472	.0717	.1372
19	.6540	.2244	.0787	.4277	.1168	.1947
20	.6566	.2397	.0933	.4322	.0818	.1678
21	.7055	.3082	.1487	.4378	.0648	.2027
22	.6013	.2144	.0866	.3615	.0661	.1305
23	.8533	.4637	.2763	.7281	.0882	.3247
24	.8317	.4413	.2632	.6917	.0539	.2453
25	.7226	.3009	.1372	.5221	.1048	.2373
26	.7457	.3182	.1424	.5561	.1070	.2463
27	.7154	.3392	.1911	.5117	.0683	.1825
28	.7372	.3526	.1967	.5434	.0582	.1935
29	.6756	.2469	.0931	.4564	.0870	.1793
30.	.5354	.1532	.0466	.2866	.0649	.1110
31	.7084	.2929	.1359	.5018	.0855	.1929
32	.5869	.1772	.0547	.3445	.0853	.1400
33	.6639	.2326	.0840	.4408	.1309	.2146
34	.6279	.2272	.0899	.3943	.0958	.1646
35	.6688	.2670	.1189	.4472	.0736	.1808
36	.7833	.3418	.1541	.6136	.1543	.2957
37	.5701	.1707	.0524	.3251	.0826	.1346
38	.6222	.2132	.0767	.3871	.0872	.1614
39	.8730	.5330	.3756	.7622	.0481	.2737
40	.5929	.1945	.0680	.3516	.0757	.1336
41	.5299	.1483	.0436	.2808	.0719	.1104
42	.6270	.2226	.0848	.3932	.0779	.1525
43	.6553	.2606	.1168	.4294	.0921	.1877
44	.6446	.2277	.0858	.4155	.0920	.1739
45	.6092	.2168	.0796	.3712	.0843	.1514
46	.7360	.3058	.1355	.5417	.1108	.2312
47	.6702	.2387	.0883	.4492	.1038	.1915
48	.7330	.3238	.1596	.5373	.1004	.2452

Table A.4

Values of the Six Invariants for Computing the First Four Cumulants of trace B for 48 Samples of Size 8 from the Normal Distribution.

Sample No	$\Sigma(a_p^2+b_p^2)^2$	$\Sigma(a_p^2+b_p^2)^3$	$\Sigma(a_p^2+b_p^2)^4$	$[\Sigma(a_p^2+b_p^2)^2]^2$	$\Sigma''(a_pa_q+b_pb_q)^4$	$\Sigma''(a_pb_q-b_pa_q)^4$
1	1.0015	.6292	.4343	1.0031	.0513	.3892
2	1.2644	.9405	.7237	1.5987	.0462	.7242
3	.6550	.2643	.1205	.4290	.0719	.1802
4	.7735	.3559	.1796	.5983	.0688	.2346
5	.6607	.2614	.1135	.4365	.0870	.2002
6	.6001	.2049	.0771	.3601	.0763	.1404
7	.6387	.2378	.0980	.4079	.0708	.1534
8	.6303	.2165	.0791	.3973	.0871	.1567
9	.6341	.2280	.0886	.4021	.0834	.1665
10	.8724	.4909	.3028	.7611	.0429	.2620
11	.8627	.4267	.2232	.7442	.0706	.2704
12	.8140	.3884	.2002	.6627	.0902	.2836
13	.6483	.2293	.0843	.4203	.0800	.1541
14	.7644	.3229	.1441	.5843	.0950	.2350
15	.7799	.3872	.2166	.6083	.0567	.2396
16	.9019	.4972	.2979	.8134	.0638	.3311
17	.8231	.4168	.2267	.6775	.0518	.2739
18	.5479	.1606	.0495	.3002	.0689	.1181
19	.9212	.5494	.3718	.8486	.0548	.2890
20	.6870	.3060	.1596	.4720	.0647	.1725
21	.7018	.2751	.1132	.4925	.0910	.2035
22	.6639	.2546	.1051	.4407	.0686	.1671
23	.8282	.4496	.2842	.6859	.0844	.2357
24	.9224	.5083	.3046	.8508	.1200	.3948
25	.8862	.4910	.2928	.7854	.0597	.3225
26	.7951	.4111	.2401	.6322	.0589	.2157
27	.8317	.4139	.2192	.6918	.0599	.2402
28	.7891	.3527	.1679	.6227	.1171	.2760
29	.7616	.3352	.1597	.5800	.0901	.2429
30	.6450	.2368	.0932	.4160	.0748	.1591
31	.9532	.5412	.3271	.9086	.1253	.4450
32	.6897	.2780	.1247	.4757	.0660	.1676
33	.9396	.5105	.2931	.8828	.1191	.4096
34	.5963	.2046	.0777	.3556	.0686	.1413
35	.6690	.2624	.1132	.4475	.0919	.1990
36	.7151	.3359	.1870	.5114	.0783	.2045
37	.6428	.2258	.0836	.4132	.0805	.1583
38	.6895	.2819	.1253	.4754	.1020	.2256
39	1.0418	.6310	.3977	1.0854	.0519	.4285
40	.8143	.4019	.2131	.6631	.0978	.3088
41	.7061	.2939	.1316	.4986	.0677	.1777
42	.8562	.4496	.2553	.7330	.0863	.3396
43	.8164	.4182	.2434	.6664	.0891	.2715
44	.6603	.2365	.0874	.4360	.0889	.1751
45	.7165	.2873	.1202	.5134	.1089	.2189
46	.7844	.3403	.1541	.6153	.1285	.2805
47	.6661	.2483	.0906	.4437	.0799	.1690
48	.8826	.4452	.2336	.7789	.0826	.3027

BIBLIOGRAPHY

1. Baker, G.A., Annals of Mathematical Statistics, 3, 1 (1932).

2. Box, G.E.P., Biometrika, 40, 318 (1953)

3. Eden, T. and Yates, F., J. Agric. Sci., 23, 6 (1933).

4. Gibbs, J. Willard, *Vector Analysis* (Edited by E.B. Wilson), Schribners (1901).

5. Hotelling, H., Annals of Mathematical Statistics, 2, 360 (1931).

6. James, G.S., Biometrika, 38, 324, (1951).

7. Pearson, E.S., Biometrika, 21, 259, (1929).

8. Rider, P.R., Annals of Mathematical Statistics, 2, 48 (1931).

9. Shewhart, W.A. and Winters, F.W., J. Am. Stat. Assoc., 23, 144 (1928)

10. Pitman, E.J.G., J. Roy. Stat. Soc., Supp. 4, 119 (1937)

11. Pitman, E.J.G., J. Roy. Stat. Soc., Supp. 4, 225 (1937)

12. Pitman, E.J.G., Biometrika, 29, 322, 1938

13. Tukey, J.W., Human Biology, 21, 65, (1949).

14. Welch, B.L., Biometrika, 29, 21 (1937)

15. Welch, B.L., Biometrika, 38, 330 (1951)

16. Welch, B.L., Biometrika, 34, 28 (1947)

17. Welch, B.L., Biometrika, 36, 293 (1949)